SIMULTANEOUS SELECTION OF EXTREME POPULATIONS : SOME OPTIMAL DECISION RULES

by

NEERAJ MISRA

DEPARTMENT OF MATHEMATICS

INDIAN INSTITUTE OF TECHNOLOGY KANPUR

DECEMBER, 1990

SIMULTANEOUS SELECTION OF EXTREME POPULATIONS : SOME OPTIMAL DECISION RULES

A Thesis Submitted

in Partial Fulfilment of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

by

NEERAJ MISRA

to the

DEPARTMENT OF MATHEMATICS

INDIAN INSTITUTE OF TECHNOLOGY KANPUR

DECEMBER, 1990

DEDICATED

to

My Mother : *Mrs. Girja Devi*

and

My Father : *Mr. ANANT RAM*

CERTIFICATE

It is certified that the work contained in the thesis entitled *"SIMULTANEOUS SELECTION OF EXTREME POPULATIONS : SOME OPTIMAL DECISION RULES"* by "Neeraj Misra" has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

Dr. I.D.Dhariyal

Department of Mathematics

I.I.T.Kanpur

December, 1990

ACKNOWLEDGEMENT

I am extremely grateful to my supervisor Dr. I.D.Dhariyal for his encouragement, supervision and excellent guidance during the preparation of this thesis.

I am also grateful to Professor R.K.S.Rathore for his encouragement during the course of my study.

I want to extend my appreciation to my friends Shailesh, Umesh, Ashu, Rajesh, Sudhir, Somesh, Smita, Kalika, Subha and Tanuja for their help and affection.

I gratefully acknowledge the research scholarship and other facilities provided by I.I.T.Kanpur.

Finally I thank Mr. G.L.Misra for typing the thesis.

Neeraj Misra

December, 1990

SYNOPSIS

The first formulation of the statistical infernce problems in the now-familiar "Ranking and Selection" framework was given by Bechhofer (1954) and Gupta (1956). Since the inception of this area of research a number of variations, modifications, and generalizations of the problem have been considered by many active researchers.

Let $X_1,\ldots,X_k$ denote k (≥ 2) random variables representing populations $\Pi_1,\ldots,\Pi_k$, respectively. Suppose that X_i has a probability density function $f_{X_i}(.;\theta_i)$, where θ_i is a real parameter, $i = 1,\ldots,k$. Let $\theta_{[1]} \leq \ldots \leq \theta_{[k]}$ denote the ranked values of $\theta_1,\ldots,\theta_k$, and let $\Pi_{(i)}$ denote the population associated with $\theta_{[i]}$, the i-th smallest of θ_i's. It is assumed that there is no prior knowledge about which of $\Pi_1,\ldots,\Pi_k$ is $\Pi_{(i)}$, $i=1,\ldots,k$ and that $\theta_1,\ldots,\theta_k$ are unknown. The populations $\Pi_{(1)}$ and $\Pi_{(k)}$ are called the lower extreme population (LEP) and the upper extreme population (UEP), respectively. Under the assumption that the parameters are location parameters ($f_{X_i}(x;\theta_i) = f(x-\theta_i)$ for some probability desity function $f(x)$, $i = 1,\ldots,k$), Mishra and Dudewicz (1987) considerd the problem of simultaneously selecting LEP and UEP and proposed selection procedures in the so-called "subset selection" formulation of Gupta (1956). Risko (1985) has studied the problem of selecting the UEP when $k = 2$ and θ_i is the single trial success probability of a binomial distribution, $i=1,2$.

The present dissertation continues the study of Mishra and Dudewicz (1987) and Risko (1985) by contributing some optimal procedures for simultaneous selection of LEP and UEP. A brief account of work reported in the thesis is presented in the following paragraphs.

In Chapter I a detailed review of the literature on ranking and selection procedures related to our study is given. In this review we have also included a few papers which deal with some optimality aspects of the selection procedures which we have not studied in the dissertation. The papers included in the review are classified into ten groups, namely, (i) Classical procedures, (ii) Best invariant and Bayes procedures, (iii) Minimax procedures, (iv) Two-stage procedures, (v) Γ-minimax procedures, (vi) Locally optimal procedures, (vii) Empirical Bayes procedures, (viii) Asymptotically consistent procedures, (ix) Optimal procedures under heteroscedasticity, and (x) Procedures for simultaneous selection of extreme populations.

In Chapter II we approach the problem of simultaneously selecting two non-empty subsets S_L and S_U, containing LEP and UEP, respectively, from Bayesian point of view. We extend some results of Goel and Rubin (1977) who studied the problem of selecting a non-empty subset containing UEP under a specific loss function. If the probability density functions possess monotone likelihood ratio property and if the prior distribution is symmertric, then (a) an essentially complete class of rules for finding Bayes rule with respect to a general loss function is obtained; (b) for a

semi-additive and non-negative loss function, Bayes rule is derived; (c) it is shown that Bayes rule of (b) is minimax and admissible; (d) for a specific loss function, an essentially complete class of rules for finding Bayes rule is obtained when the selected subsets are required to be disjoint; and (e) the results of (d) are applied to normal population models.

In Chapter III we study Bayes-P^* and minimax rules for the problem of simultaneous selection of LEP and UEP, extending the results of Berger (1979), Gupta and Yang (1985), and Gupta and Miescke (1986). Under the assumptions similar to those of Chapter II (a) an essentially complete class of rules for finding Bayes-P^* rules is determined; (b) a modification of Bayes-P^* rules is studied and Bayes rule is obtained for the modified formulation; (c) it is shown that for a specific loss function selection rule of Mishra and Dudewicz (1987) is minimax in the class of rules satisfying P^* condition; and (d) for the normal models, it is shown that the natural decision rule which selects the populations associated with the smallest and the largest observations as LEP and UEP, respectively, is minimax if and only if the population variances are equal.

Chapter IV deals with the optimal two-stage procedures for simultaneous selection of extreme populations. After formulating the problem and defining the two-stage procedures, we derive two-stage Bayes and two-stage permutation invariant procedures for a general loss function under the normal model.

Under the indifference-zone formulation, the problem of finding minimax rules for selecting the better of two binomial populations with unequal sample sizes is studied in Chapter V. When the risk is measured by the probability of incorrect selection, an attempt has been made to characterize the minimax rules. Some necessary conditions for a selection rule to be minimax are derived. Most of the results of this chapter have been published.

LIST OF CONTENTS

CHAPTER I INTRODUCTION AND A REVIEW OF LITERATURE

1.1 Introduction and Outline 1

1.2 Definitions and Notation 2

1.3 Review of the Literature 6

CHAPTER II SIMULTANEOUS SELECTION OF EXTREME POPULATIONS : A BAYESIAN APPROACH

2.1 Introduction 42

2.2 An Essentially Complete Class For Bayes Rules 44

2.3 Semi-additive And Non-negative Loss Function 53

2.4 Bayes Rules When The Selected Subsets Are Required To Be Disjoint: Specific Loss Function 58

2.5 Application To Normal Distribution 70

CHAPTER III SIMULTANEOUS SELECTION OF EXTREME POPULATIONS : BAYES - P^* AND MINIMAX RULES

3.1 Introduction 78

3.2 Bayes - P^* Rules 79

3.3 Minimax Rules : Specific Loss Function 90

3.4 Minimax Rules Under Heteroscedasticity 104

CHAPTER IV SIMULTANEOUS SELECTION OF EXTREME POPULATIONS : OPTIMAL TWO-STAGE PROCEDURES

4.1 Formulation Of Problem 114

4.2 The Optimal Second Stage Decision Rule 119

4.3 The Optimal Two-Stage Decision Rule 132

CHAPTER V SELECTING THE BETTER OF TWO BINOMIAL POPULATIONS : OPTIMAL DECISION RULES

5.1 Introduction 136

5.2 Formulation Of The Problem 137

5.3 Some Necessary Conditions For Minimaxity 139

REFERENCES 154

CHAPTER I

INTRODUCTION AND A REVIEW OF LITERATURE

1.1 Introduction and Outline

About thirty five years ago statistical inference problems were first formulated in the now-familiar "ranking and selection" framework in their pioneering work by Bechhofer (1954) and Gupta (1956). A large number of active researchers have contributed to the area of ranking and selection procedures since then. The present dissertation continues this study further by contributing some optimal procedures for selecting the extreme populations.

In order to make the presentation as self contained as possible, we give a detailed review of the relevant literature in Section 1.3 of this chapter. Some preliminary definitions and notation are introduced in Section 1.2.

The problem of finding Bayes procedures for selecting extreme populations is considered in Chapter II. For the goal where selected subsets need not be disjoint, an essentially complete class of rules for finding Bayes rule is derived for a very general loss function and results obtained are applied to more restrictive class of loss functions. The goal of simultaneously selecting disjoint subsets containing extreme populations is also considered. For a particular loss, an essentially complete class of rules for finding Bayes rule is obtained and application to normal distribution is made.

Chapter III deals with the problem of finding Bayes-P^* rules (Bayes rules for which posterior probability of simultaneously selecting the populations associated with the smallest and the largest parameters is atleast P^*, where P^* is a pre-assigned constant) and minimax rules. For normal distributions and for a very general loss function, an essentially complete class of two-stage rules for finding best invariant two-stage rule is obtained in Chapter IV.

The problem of selecting the better of two binomial populations with unequal sample sizes is considered in Chapter V. Necessary conditions for a rule to be minimax are derived.

1.2 Definitions and Notation

For uniformity we will adopt the following notation in this thesis. Let $X_1,\ldots,X_k$ be k ($\geq$ 2) random variables (r.v.'s) from populations $\pi_1,\ldots,\pi_k$, respectively and suppose that the random vector $\underset{\sim}{X} = (X_1,\ldots,X_k)$ has probability density function (pdf) $g_{\underset{\sim}{X},\underset{\sim}{\alpha}}(.;\underset{\sim}{\theta})$ depending on k-dimensional vectors of parameters $\underset{\sim}{\theta} = (\theta_1,\ldots,\theta_k)$ and $\underset{\sim}{\alpha} = (\alpha_1,\ldots,\alpha_k)$. Let $\Omega \times A$ denote the parameter space. In most of our presentation we will be dealing with the case where $X_1,\ldots,X_k$ are independent and $\underset{\sim}{\alpha}$ is known, in which case we will suppress the dependence of the joint pdf on $\underset{\sim}{\alpha}$ by dropping it and denote our parameter space by Ω, pdf of

$\underset{\sim}{X} = (X_1,\ldots,X_k)$ by $g_{\underset{\sim}{X}}(.;\underset{\sim}{\theta})$ and pdf of X_i by $f_{X_i}(.;\theta_i)$ $(i = 1,\ldots,k)$ so that $g_{\underset{\sim}{X}}(.;\underset{\sim}{\theta}) = \prod_{i=1}^{k} f_{X_i}(.;\theta_i)$.

Definition 1.2.1 Let $\theta_{[1]} \le \ldots \le \theta_{[k]}$ denote the ranked values of $\theta_1,\ldots,\theta_k$. Let $\pi_{(i)}$ denote ~~the~~ a population associated with $\theta_{[i]}$, the i-th smallest of θ_i's. Any other population or sample quantity associated with $\pi_{(i)}$ will be denoted with a subscript (i) attached to it.

Throughout we assume that there is no prior knowledge about which of $\pi_1,\ldots,\pi_k$ is $\pi_{(i)}$, $i = 1,\ldots,k$ and that $\theta_1,\ldots,\theta_k$ are unknown.

Definition 1.2.2 Let $X_{i1},\ldots,X_{in_i}$ denote a sample of size n_i from π_i, $i = 1,\ldots,k$. Let $T_i = T_i(X_{i1},\ldots,X_{in_i})$ be a function of the sample from π_i. Let $T_{[1]} \le \ldots \le T_{[k]}$ denote the ranked $T_1,\ldots,T_k$. Then we denote the population associated with $T_{[i]}$, the i-th smallest of T_i's, by $\pi_{\{i\}}$. Any other population or sample quantity associated with $\pi_{\{i\}}$ will be denoted with a subscript $\{i\}$ attached to it.

Definition 1.2.3 Populations $\pi_{(1)}$ and $\pi_{(k)}$ are called the lower extreme population (LEP) and the upper extreme population (UEP) respectively. Sometimes UEP will also be referred to as the "best" population.

Definition 1.2.4 Let $\Omega(\Delta^*) = \{\underset{\sim}{\theta} : \theta_{[k]} - \theta_{[k-1]} \ge \Delta^*\}$ and $\Omega(\Delta_1^*,\Delta_2^*) = \{\underset{\sim}{\theta} : \theta_{[2]} - \theta_{[1]} \ge \Delta_1^*,\ \theta_{[k]} - \theta_{[k-1]} \ge \Delta_2^*\}$, for some $\Delta^*,\ \Delta_1^*,\ \Delta_2^* > 0$.

Definition 1.2.5 If θ is a location (scale) parameter of the distribution of r.v. X, then $f_X(x,\theta) = f(x-\theta)$ ($\frac{1}{\theta} f(x/\theta)$) where $f(.)$ is a pdf independent of θ.

Definition 1.2.6 A pdf $f(.;\theta)$ has monotone likelihood ratio (MLR) property if for $x_1 \geq x_2$ and $\theta_1 \geq \theta_2$

$$f(x_1;\theta_1)f(x_2;\theta_2) - f(x_1;\theta_2)f(x_2;\theta_1) \geq 0.$$

Definition 1.2.7 A joint pdf $g(\underset{\sim}{x};\underset{\sim}{\theta})$ is said to have generalized MLR in $\underset{\sim}{x}$ if for any i and for fixed $x_1,\ldots,x_{i-1},x_{i+1}\ldots x_k$, $g(\underset{\sim}{x};\underset{\sim}{\theta}^*)/g(\underset{\sim}{x};\underset{\sim}{\theta}^{**})$ is nondecreasing in x_i, where $\underset{\sim}{\theta}^* = (\theta_1^*,\ldots,\theta_k^*)$, $\underset{\sim}{\theta}^{**} = (\theta_1^{**},\ldots,\theta_k^{**})$, $\theta_i^* > \theta_i^{**}$ and $\theta_j^* = \theta_j^{**}$, $j = 1,\ldots,i-1, i+1,\ldots,k$.

Definition 1.2.8 A distribution function (df) $F(.;\theta)$ is said to have stochastically increasing property (SIP) if for $\theta_1 \geq \theta_2$

$$F(x;\theta_1) \leq F(x;\theta_2) \;, \quad \forall\, x.$$

Definition 1.2.9 Let $\mathscr{A}$ denote the action space associated with a problem under study and suppose that G denotes the group of permutations on the indices $\{1,\ldots,k\}$.

In the selection problems that we consider $\mathscr{A}$ generally consists of either a subset or a pair of subsets of the set of populations $\{\pi_1,\ldots,\pi_k\}$. Any such subset $B = \{\pi_{i_1},\ldots,\pi_{i_\ell}\}$ will be written simply as $\{i_1,\ldots,i_\ell\}$ so that $j \in B$ will mean $\pi_j \in B$.

Definition 1.2.10 Let $g \in G$, $B = \{i_1,\ldots,i_\ell\} \subset \{1,\ldots,k\}$ and $\underset{\sim}{x} = (x_1,\ldots,x_k)$. Then $gB = \{gi_1,\ldots,gi_\ell\}$ denotes the image of B under g and $g\underset{\sim}{x} = (x_{g^{-1}1},\ldots,x_{g^{-1}k})$, where g^{-1} denotes the inverse

Definition 1.2.11 A loss function L is said to be permutation invariant if $L(\underset{\sim}{\theta},a) = L(g\underset{\sim}{\theta},ga)$ (or $L(\underset{\sim}{\theta},(a_1,a_2)) = L(g\underset{\sim}{\theta},(ga_1,ga_2))$ in case $\mathscr{A}$ contains pairs of subsets), $\forall\ g \in G$, $\underset{\sim}{\theta} \in \Omega$ and $a \in \mathscr{A}$ ($(a_1,a_2) \in \mathscr{A}$).

Definition 1.2.12 A decision rule δ is a measurable map from $\mathscr{X} \times \mathscr{A}$ to $[0,1]$ such that $\sum_{a\in\mathscr{A}} \delta(a|\underset{\sim}{x}) = 1$, $\forall\ \underset{\sim}{x} \in \mathscr{X}$, where $\mathscr{X}$ is the sample space of the observation vector $\underset{\sim}{X} = (X_1,\ldots,X_k)$. Let $\mathscr{D}$ denote the class all decision rules δ for a given decision problem.

Definition 1.2.13 Let δ be a decision rule for selecting UEP. Then $\psi_i^{\delta}(\underset{\sim}{x}) = \sum_{a\in\mathscr{A}:\, i\in a} \delta(a|\underset{\sim}{x})$, $i = 1,\ldots,k$. are called individual selection probabilities associated with decision rule δ.

For any decision rule δ for simultaneously selecting LEP and UEP

$$\psi_{i,j}^{\delta}(\underset{\sim}{x}) = \sum_{(a_1,a_2)\in\mathscr{A}:\, i\in a_1,\, j\in a_2} \delta((a_1,a_2)|\underset{\sim}{x}) \quad (i,j = 1,\ldots,k)$$

are referred to as pair selection probabilities associated with δ.

Definition 1.2.14 A decision rule δ for selecting UEP (simultaneously selecting LEP and UEP) is said to be permutation invariant if $\delta(a|\underset{\sim}{x}) = \delta(ga|g\underset{\sim}{x})$, $\forall$ g,a, and $\underset{\sim}{x}$ ($\delta((a_1,a_2)|\underset{\sim}{x}) = \delta((ga_1,ga_2)|g\underset{\sim}{x})$, $\forall$ g, (a_1,a_2), and $\underset{\sim}{x}$).

Definition 1.2.15 A decision rule δ for selecting UEP is said to be monotone if for every $i \in \{1,\ldots,k\}$ and $\underset{\sim}{x}$, $\underset{\sim}{x}' \in \mathscr{X}$ with $x_i \le x'_i$ and $x_j \ge x'_j$, $j \ne i$, we have $\psi_i^{\delta}(\underset{\sim}{x}') \ge \psi_i^{\delta}(\underset{\sim}{x})$.

Definition 1.2.16 Let $\phi(.)$ and $\Phi(.)$ denote the standard normal pdf and df respectively, that is

$$\phi(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}, \quad -\infty < y < \infty$$

and

$$\Phi(y) = \int_{-\infty}^{y} \phi(x)\, dx, \quad -\infty < y < \infty.$$

Definition 1.2.17 Let $F(.)$ denote some df Then for $a \neq -1$, $F^a(x)$ denotes $[F(x)]^a$, while $F^{-1}(y)$ denotes the inverse of $F(.)$ evaluated at y.

Definition 1.2.18 A prior distribution $\tau(\underset{\sim}{\theta})$ on Ω is said to be permutation invariant (symmetric) if $\tau(\underset{\sim}{\theta}) = \tau(g\underset{\sim}{\theta})$, $\forall\ \underset{\sim}{\theta} \in \Omega$ and $g \in G$.

Definition 1.2.19 A " Correct Selection " (CS) is defined to be an event which fulfils the goal of the problem at hand.

Some other definitions and notation, which are used less frequently, will be introduced as and when such need arises.

1.3 Review of the Literature

The work of Bechhofer (1954) and Gupta (1956) gave rise to two main approaches to ranking and selection problems which have come to be known as the indifference-zone approach and the subset selection approach. In indifference-zone approach the basic goal is to select certain "best" populations. The procedure should be such that the probability of achieving the goal is at least P^* whenever the parameter vector $\underset{\sim}{\theta}$ lies in a certain subset of Ω called the preference-zone. The region complementary to the

preference-zone is called the indifference-zone. In subset selection approach the goal is to select a non-empty subset S of $\{\pi_1, \ldots, \pi_k\}$ such that the selected subset contains certain "best" populations. The selection procedure should be such that it achieves the goal with probability at least equal to P^* for all $\underset{\sim}{\theta} \in \Omega$. Since the inception of the field, many variations and generalizations of the above two goals have been considered.

In this section we review the literature on ranking and selection related to our study. In this review we have also included few papers which deal with some optimality aspects of the selection procedures which we have not studied in this dissertation. The papers included in review are classified into ten groups, namely, (I) Classical procedures, (II) Best invariant and Bayes procedures, (III) Minimax procedures, (IV) Two-stage procedures, (V) Γ-minimax procedures, (VI) Locally optimal procedures, (VII) Empirical Bayes procedures, (VIII) Asymptotic consistent procedures, (IX) Optimal procedures under heteroscedasticity, and (X) Procedures for simultaneous selection of extreme populations.

(I) Classical Procedures

Consider the goal of partitioning k normal populations with means $\theta_1, \ldots, \theta_k$ into s disjoint and non-empty subsets $A_1, \ldots, A_s$ such that A_1 contains k_1 worst populations (populations associated with k_1 smallest means), A_2 contains k_2 next worst populations (populations associated with k_2 next smallest means), ..., and A_s contains k_s best populations (populations associated with k_s largest means), $k_1 + \ldots + k_s = k$. For $i=1,\ldots,k$, let $\overline{X}_i$

denote the sample mean based on a random sample of size n_i from π_i. Assuming that the variances $\sigma_1^2,\ldots,\sigma_k^2$ are known, Bechhofer (1954) proposed a single sample procedure which assigns populations associated with $\bar{X}_{[\hat{k}_{r-1}+1]}, \bar{X}_{[\hat{k}_{r-1}+2]},\ldots,\bar{X}_{[\hat{k}_r]}$ to A_r, where $\hat{k}_r = \sum_{i=1}^{r} k_i$ and $\hat{k}_0 = 0$, and showed that this procedure achieves the goal with probability at least P^* $\left(\frac{1}{\binom{k}{k_1 \ldots k_s}} < P^* < 1\right)$ whenever $\Delta_{\hat{k}_i+1,\hat{k}_i} = \theta_{[\hat{k}_i+1]} - \theta_{[\hat{k}_i]} \geq \delta^*_{\hat{k}_i+1,\hat{k}_i}$, $i = 1,\ldots,s-1$. In particular case when $s = 2$, $k_1 = k-1$, $k_2 = 1$, $n_1 = n_2 = n$, $\sigma_i^2 = \sigma^2$, $i = 1,\ldots,k$, the common sample size n required to achieve the goal is given by

$$\int_{-\infty}^{\infty} \Phi^{k-1}(y + \sqrt{n}\,\delta^*_{k,k-1}/\sigma)\,\phi(y)\,dy = P^* \qquad (1.3.1)$$

Values of n satisfying (1.3.1) are tabulated for various values of P^*, k, and $\delta^*_{k,k-1}$.

Gupta (1956) proposed decision rules for selecting non-empty subsets of k normal populations which contain the "best" population (associated with the largest mean $\theta_{[k]}$) with probability at least P^*. The variances were assumed to be equal and cases of known and unknown common variance were treated separately. Let σ^2 be the common variance and let $\bar{X}_i$ denote the mean of a random sample of size n from π_i, $i=1,\ldots,k$. In the known σ^2 case the proposed rule is

R_1 : Select π_i if and only if $\bar{X}_i \geq \bar{X}_{[k]} - \frac{d\sigma}{\sqrt{n}}$

where $d = d(k,P^*) > 0$ is chosen such that

$$\inf_{\Omega} P_{\underset{\sim}{\theta}}(\pi_{(k)} \text{ is selected}) = P^*.$$

When σ^2 is unknown the rule is same as R_1 except that σ^2 is replaced by the sample variance s^2. The expected subset size is shown to be less than or equal to kP^*.

Seal (1955, 1957) proposed a class $\mathscr{C}$ of decision rules $D(c_1,\ldots,c_{k-1})$ $(c_i \geq 0, \sum_{i=1}^{k-1} c_i = 1)$ for selecting a subset of k normal populations containing the best. His class $\mathscr{C}$ consists of decision rules of the form

$$\text{Select } \pi_i \text{ iff } \bar{X}_i \geq \sum_{j=1}^{k-1} c_j \bar{X}^{(i)}_{[j]} - st_{P^*}(c_1,\ldots,c_{k-1})$$

where $\bar{X}_i$ is the mean of the sample from π_i, $\bar{X}^{(i)}_{[1]} \leq \ldots \leq \bar{X}^{(i)}_{[k-1]}$ denote the ordered $\bar{X}_j$'s after deleting $\bar{X}_i$ $(i = 1,\ldots,k)$, and s^2 is the usual pooled unbiased estimate of common variance σ^2. Constants $c_1, \ldots, c_{k-1}$ and $t(P^*,c_1,\ldots,c_{k-1})$ are chosen so as to satisfy the probability requirement. Let D_0 denote the decision rule with $c_i = \frac{1}{k-1}$, $i=1,\ldots,k-1$ and let $D(r)$ denote the decision rule with $c_r = 1$ and $c_j = 0$, $j \neq r$, $r=1,\ldots,k-1$. A comparison between rules $D(1)$, $D(k-1)$ and D_0 is made and it is recommended that decision rule $D(k-1)$ should be used. Rules D_0 and $D(k-1)$ are referred to as average type and Gupta's maximum type rules, respectively.

Let $X_1,\ldots,X_k$ be independent random variables representing binomial populations $\pi_1,\ldots,\pi_k$ with common number of trials n and single trial success probabilities $\theta_1, \ldots, \theta_k$, respectively. For the problem of selecting the 'best' population Sobel and Huyett (1957) propose the rule which selects the population associated with $X_{[k]}$. They show that for this rule the infimum of $P_{\underset{\sim}{\theta}}(CS)$

over $\Omega(\Delta^*)$ is attained when $\theta_{[1]} = \ldots = \theta_{[k-1]} = \theta_{[k]} - \Delta^*$, where $\Delta^* \in (0,1)$ is a pre-specified constant. Tables giving smallest common sample size n needed to guarantee $P_{\underset{\sim}{\theta}}(CS) \geq P^*$, $\forall\, \underset{\sim}{\theta} \in \Omega(\Delta^*)$ for various values of k, Δ^* and P^* are provided.

For selecting a subset of k binomial populations containing the best, Gupta and Sobel (1960) propose a maximum type procedure

R_M : Select π_i if and only if $X_i \geq X_{[k]} - d(k,P^*)$

where $d(k,P^*) > 0$ is a constant chosen so that

$$P_{\underset{\sim}{\theta}}(CS) \geq P^*, \quad \forall\, \underset{\sim}{\theta} \in \Omega.$$

It is proved that $\inf_{\Omega} P_{\underset{\sim}{\theta}}(CS)$ occurs when $\theta_1 = \ldots = \theta_k$ but $\inf_{\Omega} P_{\underset{\sim}{\theta}}(CS)$ is not independent of this common value, say θ. For k=2 $\inf_{\Omega} P_{\underset{\sim}{\theta}}(CS)$ occurs at $\theta = \frac{1}{2}$ but for $k > 2$, the common value at which $\inf_{\Omega} P_{\underset{\sim}{\theta}}(CS)$ occurs is not known. However, as $n \longrightarrow \infty$, this common value tends to $\frac{1}{2}$.

Gupta (1965) proposed procedures for selecting subsets containing the populations with the largest location (scale) parameter. Let $X_1,\ldots,X_k$ be independent random variables with pdf's $f(.;\theta_1),\ldots,f(.;\theta_k)$, respectively, having the MLR property. In the location parameter case the rule is

R_2 : Select π_i iff $X_i \geq X_{[k]} - d(k,P^*)$

and in the scale parameter case the rule is

R_3 : Select π_i iff $X_i \geq c(k,P^*)\, X_{[k]}$

where $d > 0$ and $c \in (0,1)$ are chosen so that

$$\inf_{\Omega} P_{\underset{\sim}{\theta}}(CS) = P^*.$$

Let $|S|$ denote the cardinality of the subset S selected by rule R_2 (or R_3). Then it is shown that

(i) $\inf_{\Omega} P_{\underset{\sim}{\theta}}(CS)$ and $\sup_{\Omega} E_{\underset{\sim}{\theta}}(|S|)$ occurs at $\theta_1 = \ldots = \theta_k$; (ii) $P_{\underset{\sim}{\theta}}(\pi_i \in S) \geq P_{\underset{\sim}{\theta}}(\pi_j \in S)$, if $\theta_i > \theta_j$; (iii) $\sup_{\Omega} E_{\underset{\sim}{\theta}}(|S|) = kP^*$ and that $P_{\underset{\sim}{\theta}}(\pi_{(k)} \in S) \geq P_{\underset{\sim}{\theta}}(\pi_{(k)} \in S^*)$ for every subset S^* such that $|S^*|=|S|$. Tables for applications are provided.

(II) <u>Best Invariant and Bayes Procedures</u>

<u>Bahadur</u> <u>and</u> <u>Goodman</u> (1952) explored the problem of selecting UEP using decision-theoretic approach. Suppose T_i denotes sufficient statistic based on observations from π_i and suppose that joint pdf of $\underset{\sim}{T} = (T_1, \ldots, T_k)$ is given by

$$g_{\underset{\sim}{T}}(\underset{\sim}{t};\underset{\sim}{\theta}) = C(\underset{\sim}{\theta}) \prod_{i=1}^{k} f(t_i;\theta_i)$$

where $f(.;\theta_i)$ has MLR property. Assume that loss incurred in selecting π_i when $\underset{\sim}{\theta}$ is true parametric value is $L_i(\underset{\sim}{\theta})$. Under the assumption that $L_i(\underset{\sim}{\theta})$ is permutation invariant and $L_i(\underset{\sim}{\theta}) \leq L_j(\underset{\sim}{\theta})$ for $\theta_i \geq \theta_j$, it is established that natural decision rule which selects the population yielding $T_{[k]}$ uniformly minimizes the risk among all permutation invariant procedures based on $\underset{\sim}{T}$. Since here the group G of permutations is finite, it follows that natural decision rule is minimax and admissible.

Later <u>Lehmann</u> (1966) gave an alternative proof of the result proved by Bahadur and Goodman (1952).

For the goal of selecting a subset of $\{\pi_1,\ldots,\pi_k\}$ containing UEP Studden (1967) used the following loss function

$$L(\underset{\sim}{\theta},a) = \sum_{i \in a} \ell_i(\theta) + c\, I_{\{\pi_{(k)} \notin a\}}$$

where $c > 0$ is a constant, I_B denote the indicator function of a set B, and $L(\underset{\sim}{\theta},a) = L(g\underset{\sim}{\theta},ga)$ $\forall$ $\underset{\sim}{\theta}$,a and g. Let $\Omega_k = \{ \underset{\sim}{\theta} : \theta_i < \theta_k,$ $i=1,\ldots,k-1 \}$ and let

$$p_i(\underset{\sim}{x};\underset{\sim}{\theta}) = \frac{1}{\lfloor k-1} \sum_{g:g^{-1}k=i} g_{\underset{\sim}{X}}(\underset{\sim}{x};\underset{\sim}{\theta}).$$

It is shown that a decision rule δ is best invariant (it has the smallest risk among all invariant procedures) if and only if

$$\psi_k^{\delta}(\underset{\sim}{x}) = \begin{cases} 1 & \text{, if } cp_k(\underset{\sim}{x};\underset{\sim}{\theta}) > \sum_{i=1}^{k} \ell_i(\underset{\sim}{\theta})\, p_i(\underset{\sim}{x};\underset{\sim}{\theta}) \\ 0 & \text{, if } cp_k(\underset{\sim}{x};\underset{\sim}{\theta}) < \sum_{i=1}^{k} \ell_i(\underset{\sim}{\theta})\, p_i(\underset{\sim}{x};\underset{\sim}{\theta}) \end{cases}$$

for $\underset{\sim}{\theta} \in \Omega_k$ a.e. The functions $\psi_i^{\delta}(\underset{\sim}{x})$, $i \neq k$ are defined by the invariant condition. Applications to normal and exponential distributions are made.

Eaton (1967) considered the general goal of partitioning a set of populations $\{\pi_1,\ldots,\pi_k\}$ into s disjoint subsets $A_1,\ldots,A_s$ such that A_1 contains population corresponding to k_1 largest θ_i, A_2 contains population corresponding to next k_2 largest θ_i, ..., and A_s contains populations associated with k_s smallest θ_i, where $1 \le k_i < k$, $\sum_{i=1}^{s} k_i = k$. Here the action space consists of all partitions $\mathbf{A} = (A_1,\ldots,A_s)$ of $\{1,\ldots,k\}$. Let the element of group of permutations which interchanges i and j, leaving all other

members of $\{1,\ldots,k\}$ fixed be denoted by (i,j) and suppose $L(\underset{\sim}{\theta},\mathbf{A})$ denotes the loss associated with the partition $\mathbf{A}$ when $\underset{\sim}{\theta}$ is the true parameter value. For the partitions $\mathbf{A} = (A_1,\ldots,A_s)$ and $\mathbf{A}' = (A'_1,\ldots,A'_s)$ with $i \in A_\beta$, $i \in A'_{\beta+1}$, $j \in A'_\beta$, $j \in A_{\beta+1}$ (for some β, $0 \le \beta < s$) and $(i,j)\,\mathbf{A}' = \mathbf{A}$, suppose $L(\underset{\sim}{\theta},\mathbf{A}) \le L(\underset{\sim}{\theta},\mathbf{A}')$ when $\theta_i \ge \theta_j$ and suppose $L(\underset{\sim}{\theta},\mathbf{A}) = L(g\underset{\sim}{\theta},g\mathbf{A})$ $\forall$ $\underset{\sim}{\theta}$, $\mathbf{A}$ and g. Under the assumptions that the joint pdf of random observable $\underset{\sim}{X} = (X_1,\ldots,X_k)$ [X_i corresponding to π_i] has a certain property M [namely, $x_i \ge x_j$ and $\theta_i \ge \theta_j$ implies that $g_{\underset{\sim}{X}}(\underset{\sim}{x};\underset{\sim}{\theta}) \ge g_{\underset{\sim}{X}}(\underset{\sim}{x};(i,j)\underset{\sim}{\theta})$, where $(i,j)\underset{\sim}{\theta}$ is the vector $\underset{\sim}{\theta}$ with the components θ_i and θ_j interchanged] and the prior distribution τ is symmetric it is shown that natural decision rule which takes $A_1 = \{\pi_{(k-k_1+1)}, \ldots, \pi_{(k)}\}$, $A_2 = \{\pi_{(k-k_1-k_2+1)}, \ldots, \pi_{(k-k_1)}\}$, ..., and $A_s = \{\pi_{(1)}, \ldots, \pi_{(k_s)}\}$ is Bayes. This rule is further seen to be best invariant and hence minimax and admissible. Note that result proved by Eaton is a generalization of a result obtained by Bahadur and Goodman (1952) and Lehmann (1966). We make use of these results in Chapter II to determine a Bayes rule for simultaneous subset selection of LEP and UEP.

Alam (1973) noted that there may exist distributions for which posterior distributions possess SIP but they do not have property M. Motivated by this he studied the problem considered by Eaton (1967) for s = 2, when the posterior distribution with respect to a symmetric prior τ possesses SIP and is invariant. Under the assumption that the loss function is permutation invariant and it is non-increasing (non-decreasing) in θ_i for

$i \in A_1$ (A_2), he proved that the natural decision rule is Bayes with respect to τ. Under an additional assumption that the loss function is bounded and continuous in $\underset{\sim}{\theta}$, natural decision rule is shown to be admissible.

For selecting a subset of $\{\pi_1, \ldots, \pi_k\}$ containing UEP, Deely and Gupta (1968) considered the loss function

$$L(\underset{\sim}{\theta}, a) = \sum_{i \in a} \beta_i(a)\ (\theta_{[k]} - \theta_i), \quad \beta_i(a) \geq 0$$

Let $X_1, \ldots, X_k$ be independently distributed, X_i from π_i having a normal distribution with mean θ_i and variance 1. Suppose that prior distribution $\tau(\underset{\sim}{\theta}) = \prod_{i=1}^{k} \tau_i(\theta_i)$. Under the assumption that $\beta_i(a) = \beta > 0$ for every a whose cardinality is one, and $\sum_{i \in a} \beta_i(a) \geq \beta$ for every $a \in \mathscr{A}$, it is shown that Bayes rule selects only one population.

Chernoff and Yahav (1977) considered the goal of selecting a subset of k normal populations having means $\theta_1, \ldots, \theta_k$ and a common known variance σ^2. For the loss function

$$L(\underset{\sim}{\theta}, a) = c(\theta_{[k]} - \max_{i \in a} \theta_i) + \theta_{[k]} - \frac{1}{|a|} \sum_{i \in a} \theta_i$$

where c is a positive constant. Under the assumption that $\underset{\sim}{\theta}$ has symmetric multivariate normal prior, a Bayes procedure is obtained and compared with the procedures of Gupta (1965) and Desu and Sobel (1968). Using Monte Carlo simulations, it is empirically shown that Gupta's rule is highly efficient.

With the nonlinear loss function

$$L(\underset{\sim}{\theta}, a) = c|a| + (\theta_{[k]} - \max_{i \in a} \theta_i), \qquad c > 0,$$

Goel and Rubin (1977) investigated the problem of finding Bayes rules for selecting a subset containing the best of k independent populations with pdfs having MLR property. For $j = 1,\ldots,k$, let δ_j^* be the rule which selects with probability one (w.p. 1) the subset corresponding to j largest observations and suppose that τ is a symmetric prior on Ω. It is shown that the class of rules $\{\delta_j^*, j = 1,\ldots,k\}$ is an essentially complete class for finding Bayes rules. Under an additional assumption that the prior distribution τ is a mixture of independent and identically distributed random variables, Bayes rule is simplified. An "approximate" Bayes rule is also obtained which selects larger subsets than the Bayes rule but is the Bayes rule for $k = 2$. Under the assumption that prior distribution of $\theta_{[1]}$ belongs to a location parameter family, it is observed that the approximate Bayes rule is similar to the rule proposed by Gupta (1965). In the special case, when observation X_i from π_i has normal distribution with mean θ_i and known variance σ^2 and $\underset{\sim}{\theta}$ has symmetric multivariate normal prior distribution, it is shown that if

$$\int_{-\infty}^{\infty} \Phi^m(z)\, \Phi(-z)dz < \frac{c}{\gamma} < .56419$$

then the maximum size of the selected subset is m and if $\frac{c}{\gamma} \geq .56419$, then the Bayes rule selects $\{\pi_{(k)}\}$ w.p. 1, where γ^2 is the common posterior variance. We extend these results to the problem of simultaneously selecting LEP and UEP in Chapter II.

For selecting a subset of normal populations, with unknown means and common known variance, containing the UEP Gupta and Hsu (1978) assumed that the loss function is given by

$$L(\underset{\sim}{\theta},a) = c_1 \, I_{\{\pi_{(k)} \notin a\}} + c_2 \, |a|, \qquad c_1, c_2 \geq 0, \; c_1+c_2 = 1$$

and that the prior distribution of $\underset{\sim}{\theta}$ is multivariate normal with mean vector $\mu\mathbf{1}$ and covariance matrix $rI + t\,\mathbf{1}\mathbf{1}'$, where $\mathbf{1}' = (1,\dots,1)$, $r > 0$, $-\frac{r}{k} < t$ and I denotes the identity matrix, and obtained the Bayes rule

R_B : Select π_i iff $X_i \geq X_{[k-1]}$ and/or $P(\theta_i = \theta_{[k]}|\underset{\sim}{X}) \geq c_2/c_1$.

Monte Carlo comparison of rules R_B, Gupta's (1965) maximum type and Seal's (1955) average type for k = 3 and k = 8, indicate that the maximum type procedure does almost as well as Bayes procedure.

Miescke (1979) extended the result of Deely and Gupta (1968), by assuming a loss function

$$L(\underset{\sim}{\theta},a) = \sum_{i\in a} \beta_i(a)\, \ell_i(\underset{\sim}{\theta}),$$

where $\beta_i(a) = \beta(|a|)$ satisfy $m\,\beta(m) \geq \beta(1)$, for $m = 1,\dots,k$ and ℓ_i's are non-negative. It is proved that there exists a Bayes rule which always selects only one population. For an additive loss function, $L(\underset{\sim}{\theta},a) = \sum_{i\in a} \ell_i(\underset{\sim}{\theta})$), sufficient conditions for a rule δ to be ordered [for which $\psi_j^{\delta}(\underset{\sim}{x}) \geq \psi_i^{\delta}(\underset{\sim}{x})$ whenever $x_i < x_j$, $\forall\, \underset{\sim}{x} \in \mathcal{X}$] and monotone are derived. Also for selecting a subset of normal populations, it is shown that (i) Gupta's (1965) maximum type rule is a limit of Bayes rules with respect to symmetric priors and loss function

$$L(\underset{\sim}{\theta},a) = \sum_{i\in a} \ell(\theta_{[k]} - \theta_i - \varepsilon),$$

where ℓ is bounded and increasing and (ii) Seal's (1955) average type procedure is Bayes for the unrealistic loss function

$$L(\underset{\sim}{\theta},a) = \sum_{i\in a} \left(\frac{1}{k}\sum_{j=1}^{k}\theta_j - \theta_i - \varepsilon\right), \quad \varepsilon > 0$$

Bjornstad (1981a) derived best invariant rule for the problem of selecting a subset S, containing the best, of k populations having pdfs with the MLR property, and such that $|S| \geq t$, a given number. For the loss function

$$L(\underset{\sim}{\theta},a) = \sum_{i\in a} \beta_i(a)\, \ell_i(\underset{\sim}{\theta}),$$

where $\beta_i(a) = \beta(|a|)$ satisfy $[\frac{r}{t}]\ \beta(r) \geq \beta(t)$, $\ell_i(\underset{\sim}{\theta}) \geq 0$, $\ell_i(\underset{\sim}{\theta}) = \ell_{gi}(g\underset{\sim}{\theta})$, $\forall\ \underset{\sim}{\theta} \in \Omega$, $g \in G$, and $\ell_i(\underset{\sim}{\theta}) \geq \ell_j(\underset{\sim}{\theta})$, for $\theta_i \leq \theta_j$, it is proved that natural decision rule which selects the subset corresponding to the t largest observations (with ties broken at random) is best invariant. For the loss function $L(\underset{\sim}{\theta},a) = \sum_{i\in a} \ell_i(\underset{\sim}{\theta})$ with $\ell_i(\underset{\sim}{\theta})$ satisfying $\ell_i(\underset{\sim}{\theta}) \geq \ell_j(\underset{\sim}{\theta})$ for $\theta_i < \theta_j$, a class of likelihood ratio type procedures is shown to be admissible.

The concept of a Bayes-P^* decision rule was introduced by Gupta and Yang (1985). A decision rule δ is said to be Bayes-P^* if it has the smallest Bayes risk among decision rules for which posterior probability of correct selection is greater than or equal to P^*, a specified constant. With a symmetric prior $\tau(\underset{\sim}{\theta})$ on Ω define statistics $T_i(\underset{\sim}{X}) = P(\theta_i = \theta_{[k]} \mid \underset{\sim}{X})$, $i=1,\ldots,k$. A decision rule δ for the problem of selecting a subset of $\{\pi_1,\ldots,\pi_k\}$ containing UEP is said to satisfy the PP^*-condition (posterior-P^*-condition) if, $\psi^{\delta}_{\{k\}}(\underset{\sim}{x}) = P\left[\pi_{\{k\}} \text{ is selected} \middle| \delta,\ \underset{\sim}{X}=\underset{\sim}{x}\right] = 1$, and $P\left[CS \middle| \delta, \underset{\sim}{X}=\underset{\sim}{x}\right] = P\left(\text{Select } \pi_{(k)} \middle| \delta, \underset{\sim}{X}=\underset{\sim}{x}\right) \geq P^*$, for all $\underset{\sim}{x}$.

Define rules δ^B and δ^{B^*} as follows

$$\psi^{\delta^B}_{\{i\}}(\underset{\sim}{x}) = \begin{cases} 1 & , \text{ if } i \geq j \\ 0 & , \text{ otherwise} \end{cases}$$

where $1 \leq j \leq k$ is the largest integer such that $\sum_{l=j}^{k} T_{[l]}(\underset{\sim}{x}) \geq P^*$, and

$$\psi^{\delta^{B^*}}_{\{j\}}(\underset{\sim}{x}) = \begin{cases} 1, & \text{if } \sum_{l=j}^{k} T_{[l]} \leq P^*, \quad j \neq k \\ \gamma, & \text{if } \sum_{l=j+1}^{k} T_{[l]} < P^* \text{ and } \sum_{l=j}^{k} T_{[l]}(\underset{\sim}{x}) > P^*, \\ 0, & \text{otherwise} \end{cases}$$

where the constant γ is determined so that $\gamma \; T_{[j]}(\underset{\sim}{x}) + \sum_{l=j+1}^{k} T_{[l]}(\underset{\sim}{x}) = P^*$. Suppose that the pdf's under consideration have MLR property and the loss function satisfies (i) $L(\underset{\sim}{\theta}, a) = L(g\underset{\sim}{\theta}, ga)$, (ii) $L(\underset{\sim}{\theta}, a)$ is non-increasing in θ_i for $i \in a$, and (iii) $L(\underset{\sim}{\theta}, a) \leq L(\underset{\sim}{\theta}, a')$, if $a \subset a'$, for all $g \in G$, $a \in \mathscr{A}$ and $\underset{\sim}{\theta} \in \Omega$. Then, it is shown that $\delta^B(\delta^{B^*})$ is Bayes in the class $\mathbf{D}^*(\mathcal{D}^*)$ of all non-randomized (randomized) decision rules satisfying PP^*-condition. The decision rule $\delta^B(\delta^{B^*})$ is shown to be most efficient in the sense that for any decision rule δ in $\mathbf{D}^*(\mathcal{D}^*)$

$$\text{eff}(\delta | \underset{\sim}{x}) = \frac{P(\text{Selecting UEP} | \delta, \underset{\sim}{x})}{E(\text{Size of selected subset } | \delta, \underset{\sim}{x})}$$

is not greater than $\text{eff}(\delta^B | \underset{\sim}{x})$ $(\text{eff}(\delta^{B^*} | \underset{\sim}{x}))$. Further, under

PP^*-condition, the subset selected by δ^B or δ^{B^*} is always smaller than that selected by maximum type procedure proposed by Gupta (1965); for k = 2 in normal location model, the rule ψ^B is same as the maximum type rule; and for non-informative prior $\tau(\underset{\sim}{\theta})$ = constant, Gupta's (1965) rule satisfies PP^*-condition. We derive Bayes-P^* decision rules for simultaneously selecting LEP and UEP in Chapter III. Dunnett (1960) proposed a procedure for selecting the variate associated with the largest component of the mean vector of a k-variate normal distribution with known equal variances and covariances and studied its properties assigning a k-variate normal prior distribution.

(III) **Minimax Procedures**

For selecting a subset containing the UEP, Berger (1979) proved that if the loss is measured by the subset size and if one restricts to the class $\mathcal{D}_{P^*}$ of decision rules satisfying P^*-condition ($P_{\underset{\sim}{\theta}}(CS) \geq P^*$, $\forall\ \underset{\sim}{\theta} \in \Omega$) then the minimax value of the problem is kP^*. This implies that the rules R_2 and R_3 proposed by Gupta (1965) are minimax. Some necessary conditions for a rule to be minimax are derived and it is observed that if a decision rule $\delta \in \mathcal{D}_{P^*}$ is minimax for loss measured by subset size, then it is also minimax for the loss measured by number of non-best populations selected. In Chapter III we derive minimax rules in the class of rules satisfying P^*-condition for the problem of simultaneously selecting LEP and UEP.

Berger and Gupta (1980) obtained minimax decision rules in the class of non-randomized, just and translation invariant decision rules which satisfy P^*-condition, when the risk is measured by the maximum probability of including a non-best population. Under location (scale) parameter model, it is shown that decision rule $R_2(R_3)$ proposed by Gupta (1965) is minimax and

admissible in the class of non-randomized, just and translation (scale) invariant decision rules which satisfy P^*- condition. Selection in terms of k normal populations with means $\theta_1,\ldots,\theta_k$ and unequal variances $\sigma_1^2,\ldots,\sigma_k^2$ is also considered. The variances may be of the form $\sigma_i^2 = \frac{\gamma_i^2}{n_i}$, $i = 1,\ldots,k$, so that this formulation also includes unequal sample size problem. It is proved that the decision rule

$$R_{GW}: \text{Select } \pi_i \text{ iff } X_i \geq \max_{1\leq j\leq k} \left(X_j - d\left(\frac{\gamma_i^2}{n_i} + \frac{\gamma_j^2}{n_j}\right)^{\frac{1}{2}}\right)$$

proposed by Gupta and Wong (1982) is minimax in the class of non-randomized, just and translation invariant decision rules which satisfy P^*-condition.

<u>Gupta and Huang</u> (1980a) explored the problem of finding a minimax rule within the class of decision rules for which the infimum of the probability of correct selection over a subset of parameter space is guaranteed to be a specified number P^*. Let $\xi_{ij} = \xi_{ij}(\underset{\sim}{\theta})$ be some measure of separation between π_i and π_j and suppose that there exists a monotonically non-increasing function h such that $\xi_{ji} = h(\xi_{ij})$. Define $\xi^* = \max_{1\leq i\leq k} \min_{j\neq i} \xi_{ij}$, $\Omega_i = \{\underset{\sim}{\theta} : \xi_{ij} \geq \Delta^* \ \forall\ j \neq i\}$, $1 \leq i \leq k$ and $\bar{\Omega} = \bigcup_{i=1}^{k} \Omega_i$, where Δ^* and ξ_{ii} are known and $\Delta^* > \xi_{ii}$. The population associated with ξ^* is defined to be the best population. Let the statistics Z_{ij} be based on n_i and n_j independent observations from π_i and π_j $(1 \leq i, j \leq k)$, respectively and suppose that for any i, the statistic $\underset{\sim}{Z}_i = (Z_{i1},\ldots,Z_{ik})$ is sufficient and invariant under a transformation

group and $\underset{\sim}{\xi}' = (\xi_{i1},\ldots,\xi_{ik})$ is a maximal invariant under the induced group. Let $g_{\underset{\sim}{Z}_{ij}}(.;\underset{\sim}{\theta})$ be the joint pdf of Z_{ij}'s, $j \neq i$ which will be denoted by $g_0(.)$ when $\xi_{i1} = \ldots = \xi_{ik} = \xi_{ii}$ = constant and by $g_1(.)$ when $\xi_{i1} = \ldots = \xi_{ik} = \Delta^*$, $i = 1,\ldots,k$. Consider the decision rule δ^0 defined by

$$\psi_i^{\delta^0}(\underset{\sim}{z}_i) = \begin{cases} 1, & \text{if } g_1(\underset{\sim}{z}_i) > c\, g_0(\underset{\sim}{z}_i) \\ \lambda_i, & \text{if } g_1(\underset{\sim}{z}_i) = c\, g_0(\underset{\sim}{z}_i) \\ 0, & \text{if } g_1(\underset{\sim}{z}_i) < c\, g_0(\underset{\sim}{z}_i) \end{cases}$$

where c (> 0) and λ_i are determined by $\int \psi_i^{\delta^0}(\underset{\sim}{z}_i)\, g_1(\underset{\sim}{z}_i)\, d\underset{\sim}{z}_i = P^*$, $1 \leq i \leq k$. Further assume that $\dfrac{g_1(\underset{\sim}{z}_i)}{g_0(\underset{\sim}{z}_i)}$ is non-decreasing in each component of $\underset{\sim}{z}_i$, $g_{\underset{\sim}{Z}_{ij}}(.;\underset{\sim}{\theta})$ has SIP and supremum of expected size of subset selected by δ^0 is attained at $\xi_{ij} = \xi_{ii}$ = constant, for all i,j. Then for the loss measured by subset size, the decision rule δ^0 is proved to be minimax among the rules for which $P_{\underset{\sim}{\theta}}(CS) \geq P^*$, $\forall\, \underset{\sim}{\theta} \in \bar{\Omega}$. Application to normal distribution is made in which case decision rule δ^0 reduces to Seal's (1955) average type decision rule.

Now, suppose that ξ_{ii} is known and $\xi_i = \min_{j \neq i} \xi_{ij}$, $\Omega_i' = \{\underset{\sim}{\theta} : \xi_i \geq \xi_{ii}\}$, $1 \leq i \leq k$, $\xi_{ij} \geq \xi_{ii}$, $j \neq i$ in case all θ_i's are equal, and $\Omega = \bigcup_{i=1}^{k} \Omega_i'$ holds. A population π_i is called best if $\underset{\sim}{\theta} \in \Omega_i'$. Under the above set-up <u>Gupta</u> <u>and</u> <u>Huang</u> (1980b) consider the problem of finding a decision rule δ^* which controls the expected

size of selected subset at the points of the form $(\theta,\ldots,\theta)$ and which maximizes the minimum probability of correct selection. Only those decision rules are considered which depend on observations through sufficient statistic for $\underset{\sim}{\theta}$. It is proved that the class of monotone procedures forms an essentially complete class provided underlying distributions have generalized monotone likelihood ratio property.

The problem of finding a decision rule which maximizes the minimum probability of correct selection while controlling the expected size of selected subset below a specified number was also investigated by Hsu (1985). Some optimal procedures are derived and application to normal distribution is made.

(IV) Two-Stage Procedures

Suppose we have k independent normal populations $\pi_1,\ldots,\pi_k$ with means $\theta_1,\ldots,\theta_k$ and variances $\sigma_1^2,\ldots,\sigma_k^2$, respectively. Dudewicz (1971) proved that there does not exist any single-stage selection procedure whose P(CS) is independent of σ_i^2, $i = 1,\ldots,k$. In such situations one uses sequential or two-stage procedures to solve the problem.

Bechhofer, Dunnett, and Sobel (1954) used indifference-zone approach and proposed the following two-stage procedure for selecting UEP among k normal populations with variances $\sigma_i^2 = c_i\sigma^2$, where $c_i (1 \le i \le k)$ are known positive constants and σ^2 is unknown. At first stage $c_i n_o$ observations are taken from each π_i $(1 \le i \le k)$, where n_o is an integer satisfying $n = n_o \sum_{i=1}^{k} c_i - k > 0$.

Suppose

$$s_o^2 = \frac{1}{n} \sum_{i=1}^{k} \frac{1}{c_i} \sum_{j=1}^{c_i n_o} \left(X_{ij} - \frac{1}{c_i n_o} \sum_{j=1}^{c_i n_o} X_{ij}\right)^2$$

and

$$N = \max \{n_o, \left[\frac{2s_o^2 h^2}{\Delta^{*2}}\right]\},$$

where [y] denotes the smallest integer $\geq$ y. At second stage $(N-n_o)c_i$ additional observations are taken from π_i, i = 1,....,k. Then select the population associated with

$$\overline{X}_{[k]} = \max_{1 \leq i \leq k} \left(\frac{1}{c_i N} \sum_{j=1}^{c_i N} X_{ij}\right)$$

as the UEP. The constant $h = h(n,k,P^*)$ is chosen so that $P_{\underset{\sim}{\theta}}(CS) \geq P^*$, $\forall\, \underset{\sim}{\theta} \in \Omega(\Delta^*)$.

For the goal of selecting UEP using indifference-zone approach when the variances are unequal and unknown, Dudewicz and Dalal (1975) proposed a two-stage procedure which takes samples of size n_o from π_i, i = 1,....,k at stages 1. Let s_i^2 denote the sample variance based on these n_o observations from π_i and suppose h^* is the unique solution of

$$\int_{-\infty}^{\infty} F_{n_o-1}^{k-1}(z+h^*)\, f_{n_o-1}(z)\, dz = P^*,$$

where $F_{n_o-1}(.)$ and $f_{n_o-1}(.)$ denote the df and the pdf respectively of Student's-t random variable with n_o-1 degrees of freedom. At stage 2 n_i-n_o additional observations are taken from π_i, where

$$n_i = \max \left\{n_o+1, \left[\left(\frac{s_i h^*}{\Delta^*}\right)^2\right] + 1\right\}.$$

Let c_{ij}'s be constants satisfying

$$\sum_{j=1}^{n_1} c_{ij} = 1,\ c_{i1} = \dots = c_{in_0},\ \text{and}\ s_i^2 \sum_{j=1}^{n_1} c_{ij} = \left(\frac{\Delta^*}{h^*}\right)^2.$$

The proposed procedure selects as the UEP the population which yields the largest $\tilde{X}_i$, where $\tilde{X}_i = \sum_{j=1}^{n_1} c_{ij} X_{ij}$. It is shown that the probability of correct selection for this procedure is independent of variances and its infimum over $\Omega(\Delta^*)$ is attained when $\theta_{[1]} = \dots = \theta_{[k-1]} = \theta_{[k]} - \Delta^*$. Tables needed to implement the proposed procedure are given.

Tamhane and Bechhofer (1977) proposed following two-stage procedure for selecting UEP.

R_{TB}: At stage 1, take n_1 independent observations from each of k populations and select only those π_i for which $\bar{X}_i \geq \bar{X}_{[k]}-d$, $i=1,\dots,k$, where $d \geq 0$ is chosen so that $\inf_{\underset{\sim}{\theta}\in\Omega(\Delta^*)} P_{\underset{\sim}{\theta}}(CS) \geq P^*$. If only one population is selected at stage 1, then stop sampling and select this population as UEP, otherwise take n_2 additional observations from each π_i selected at stage 1 and select the population yielding largest pooled sample mean as UEP. Note that $d = 0$ gives the single-sample procedure of Bechhofer (1954). It is required that for rule R_{TB}

$$P_{\underset{\sim}{\theta}}(CS) \geq P^*, \qquad \forall\ \underset{\sim}{\theta} \in \Omega(\Delta^*) \qquad (1.3.2)$$

It is shown that there are an infinite number of combinations of (n_1, n_2, d) which satisfy (1.3.2) for any k and (Δ^*, P^*), and that different design criteria lead to different choices. They use the design criterion proposed by Alam (1970) : For given k and

specified (Δ^*, P^*), choose (n_1, n_2, d) to minimize $\sup_{\underset{\sim}{\theta} \in \Omega(\Delta^*)} E_{\underset{\sim}{\theta}}(kn_1 + Vn_2)$ subject to $\inf_{\underset{\sim}{\theta} \in \Omega(\Delta^*)} P_{\underset{\sim}{\theta}}(CS|R_{TB}) \geq P^*$, where V is a random variable which takes value zero if only one population is selected at stage 1 and is equal to cardinality of the subset selected at stage 1, if more than one population are selected at stage 1. For k = 2, the least favorable configuration (LFC) is shown to be $\theta_{[1]} = \theta_{[2]} - \Delta^*$. It was conjectured that for k > 2 also $\theta_{[1]} = \ldots = \theta_{[k-1]} = \theta_{[k]} - \Delta^*$ is the LFC. Later Miescke and Sehr (1980) gave a non-standard proof of the conjecture in case of k = 3 populations. Recently, Bhandari and Chaudhuri (1990) have shown that the conjecture concerning LFC of R_{TB} is true for all k. For k > 2, Tamhane and Bechhofer (1977) obtained a conservative solution by taking the infimum over $\Omega(\Delta^*)$ of a lower bound of $P_{\underset{\sim}{\theta}}(CS|R_{TB})$, and later Tamhane and Bechhofer (1979) improved upon it.

For the normal means with common unknown variance σ^2, Gupta and Kim (1984) proposed a two-stage elimination type procedure.

Now suppose that the same number of observations are drawn at stage 1 and stage 2 and suppose

$$S_1(\underset{\sim}{x}) = \{ i : \overline{X}_i \geq \overline{X}_{[k]} - c,\ c > 0 \text{ fixed} \}$$

$$S_2(\underset{\sim}{x}) = \{ i : \overline{X}_i \text{ is one of the t largest values of of } \overline{X}_1, \ldots, \overline{X}_k,\ t \in \{2, \ldots, k-1\} \text{ fixed} \}$$

$$S_3(\underset{\sim}{x}) = \{ i : \overline{X}_i \geq c_i,\ c_1, \ldots, c_k \in \mathbb{R}^1 \}$$

and let $d_{1,S}$ and $d_{2,S}$ be decisions at stage 2 such that

$$d_{1,S}(\underset{\sim}{X}, \underset{\sim}{Y}) = 1, \quad \text{iff } \overline{Y}_i = \max_{j \in S} \overline{Y}_j$$

$$d_{2,S}(\underset{\sim}{X},\underset{\sim}{Y}) = 1 \ , \quad \text{iff } \overline{X}_i+\overline{Y}_i = \max_{j\in S} (\overline{X}_j+\overline{Y}_j)$$

where $\underset{\sim}{Y}$ denotes the sample drawn at stage 2. Note that rule (S_1,d_{2,S_1}) corresponds to that of Tamhane and Bechhofer (1977). Gupta and Miescke (1982) prove that for (S_i,d_{j,S_i}), $i = 2,3$, $j = 1,2$, the LFC in $\Omega(\Delta^*)$ is given by $\theta_{[1]} = \ldots = \theta_{[k-1]} = \theta_{[k]}-\Delta^*$. For (S_1,d_{1,S_1}), they derive a lower bound for $P_{\underset{\sim}{\theta}}(CS)$ which is minimized in $\Omega(\Delta^*)$ at $\theta_{[1]} = \ldots = \theta_{[k-1]} = \theta_{[k]} - \Delta^*$, a result similar to that of Tamhane and Bechhofer (1977).

Gupta and Miescke (1983) define the class of two-stage invariant randomized procedures (ν,η,δ), where at stage 1, ν and η, respectively decide how many and which populations to be selected and where at stage 2 after taking additional samples from the populations selected at stage 1, δ makes a final decision for selecting UEP. Suppose that the observations from π_i have pdf $f_{X_i}(x;\theta_i) = g(\theta_i)b(x) \exp(\theta_i x)$ and let η^* and δ^* denote the decision rules which select populations corresponding to largest sufficient statistic. Let $L(\underset{\sim}{\theta},a,i)$ denote the loss incurred if subset a is selected at stage 1 and population π_i, $i \in a$ is selected at stage 2. Assume that $L(\underset{\sim}{\theta},a,i) = L(g\underset{\sim}{\theta},ga,gi)$, $\forall\ \underset{\sim}{\theta} \in \Omega$, $g \in G$, $a \in \mathscr{A}$ and $i \in a$. Further assume that for every fixed $\underset{\sim}{\theta} \in \Omega$ with $\theta_1 \le \theta_2$ and $\tilde{a} \subset \{3,\ldots,k\}$ with $0 \le |\tilde{a}| \le k-2$, the following four conditions are satisfied

(i) $\quad L(\underset{\sim}{\theta},\{2\},2) \le L(\underset{\sim}{\theta},\{1\},1)$

(ii) $\quad L(\underset{\sim}{\theta},\tilde{a} \cup \{2\},i) \le L(\underset{\sim}{\theta},\tilde{a} \cup \{1\},i)$, $\quad i \in \tilde{a}$

(iii) $L(\underset{\sim}{\theta}, \underset{\sim}{\tilde{a}} \cup \{1,2\}, 2) \le L(\underset{\sim}{\theta}, \underset{\sim}{\tilde{a}} \cup \{1,2\}, 1)$

(iv) $L(\underset{\sim}{\theta}, \underset{\sim}{\tilde{a}} \cup \{2\}, 2) \le L(\underset{\sim}{\theta}, \underset{\sim}{\tilde{a}} \cup \{1\}, 1)$, $\quad |\tilde{a}| \ge 1$

Under the additional condition that underlying pdf's are unimodal, it is shown that procedures of the type (ν, η^*, δ^*) form an essentially complete class in the class of invariant decision rules and since the group of permutations is finite, best invariant decision rule is also minimax.

The problem of finding two-stage Bayes procedures for selecting the UEP among normal populations with common known variance is studied by Gupta and Miescke (1984a). Following loss function is considered

$$L(\underset{\sim}{\theta}, a, i) = c_1 n_1 + h(\underset{\sim}{\theta}) - \theta_i + c_2 n_2 V$$

where c_1, c_2 are constants, n_1 (n_2) are the number of observations per population at stage 1 (stage 2), V is a random variable which takes value zero if only one population is selected at stage 1 and is equal to number of elements in selected subset if more than one population are selected at stage 1, and $h(\underset{\sim}{\theta})$ is any fixed real valued function. Under the above set-up, a two-stage Bayes procedure with respect to independently and identically distributed normal priors is derived. Several properties of Bayes procedures are discussed.

(V) Γ-Minimax Procedures

In Γ-minimax criterion, it is assumed that our prior information consists of a class Γ of distributions over Ω and it is desired to find a decision rule that minimizes the maximum expected risk over Γ. Note that if Γ consists of a single prior,

we have a usual Bayes rule with respect to that prior while if Γ consists of all possible priors then Γ-minimax rule is the usual minimax rule.

Gupta and Huang (1977) derived a Γ-minimax procedure for the problem of selecting a subset containing the best population. Let ξ_{ij} be some measure of separation between π_i and π_j, $\xi_i = \max_{\substack{1\le j\le k \\ j\ne i}} \xi_{ij}$ and $\Omega_i = \{\underset{\sim}{\theta} \in \Omega : \xi_i \le \xi_0\}$, $1 \le i \le k$, where ξ_0 is a given constant. The parameter space Ω is partitioned into $k+1$ mutually exclusive subsets $\Omega_0, \Omega_1, \ldots, \Omega_k$, where Ω_0 is an indifference-zone ($\Omega_0 = \phi$ is permitted). The population associated with $\min_{1\le j\le k} \xi_j$ is called the "best" population. Suppose that the loss incurred in using a decision rule δ is given by

$$L(\underset{\sim}{\theta}, \delta(\underset{\sim}{x})) = \sum_{i=1}^{k} \sum_{j=1}^{k} L_1^{(i)}(\underset{\sim}{\theta}, \psi_j^{\delta}(\underset{\sim}{x}))$$

where for any i, $1 \le i \le k$, and $\underset{\sim}{\theta} \in \Omega_i$, $L_1^{(l)}(\underset{\sim}{\theta}, \psi_j^{\delta}(\underset{\sim}{x})) = 0$ for all $l \ne i$,

$$L_1^{(i)}(\underset{\sim}{\theta}, \psi_j^{\delta}(\underset{\sim}{x})) = \begin{cases} c_{ij}\ \psi_j^{\delta}(\underset{\sim}{x}) & , \text{ for } j \ne i \\ 0 & , \text{ for } j = i, \end{cases}$$

$c_{ij} (j \ne i)$ are given positive numbers and $c_{ii} = 0$. For $\underset{\sim}{\theta} \in \Omega_0$, the loss is assumed to be zero. Further suppose $\Gamma = \{\tau(\underset{\sim}{\theta}) : \int_{\Omega_i} d\tau(\underset{\sim}{\theta}) = q_i$, $i = 1, \ldots, k\}$, so that $\sum_{i=1}^{k} q_i \le 1$, and for $i = 1, \ldots, k$, let $\underset{\sim}{\theta}_i^* \in \Omega_i$ be such that

$$\sup_{\underset{\sim}{\theta}\in\Omega_i} E_{\underset{\sim}{\theta}}\left[\sum_{j=1}^{k} c_{ij}\, \psi_j^{\delta^0}(\underset{\sim}{X})\right] = E_{\underset{\sim}{\theta}_i^*}\left[\sum_{j=1}^{k} c_{ij}\, \psi_j^{\delta^0}(\underset{\sim}{X})\right],$$

where

$$\psi_i^{\delta^0}(\underset{\sim}{x}) = \begin{cases} 1 & , \text{ if } T_i(\underset{\sim}{x}) < \min\limits_{\substack{1\le j\le k\\ j\ne i}} T_j(\underset{\sim}{x}) \\ \alpha_i & , \text{ if } T_i(\underset{\sim}{x}) = \min\limits_{\substack{1\le j\le k\\ j\ne i}} T_j(\underset{\sim}{x}) \\ 0 & , \text{ otherwise} \end{cases}$$

and $T_i(\underset{\sim}{x}) = \sum_{j=1}^{k} c_{ji}\, q_j\, f_{\underset{\sim}{X}}(\underset{\sim}{x},\theta_j^*)$, $1 \le i \le k$, $\sum_{i=1}^{k} \alpha_i = 1$. It is proved that the decision rule δ^0 is Γ-minimax and applications to different selection problems are made.

Gupta and Kim (1980) investigated Γ-minimax rules for comparing populations $\pi_1,\ldots,\pi_k$ with a control population π_0. Suppose that the pdf associated with i-th population π_i is symmetric, unimodal and has location parameter θ_i, $i = 0,1,\ldots,k$. Define a population π_i to be superior, equivalent on inferior to π_0 according as $\theta_i \ge \theta_0 + \Delta^*$, or $\theta_0 - \Delta^* < \theta_i < \theta_0 + \Delta^*$ or $\theta_i \le \theta_0 + \Delta^*$, where $\Delta^* > 0$ is specified. Only those decision rules are considered for which the selection or the rejection of the i-th population depends on observations only through X_0 and X_i With a reasonable loss function, Γ-minimax and minimax rules are derived.

Randles and Hollander (1971) and Gupta and Hsiao (1981) have developed Γ-minimax procedures for selecting all "good" populations and rejecting "bad" ones, where a population is defined to be good (bad) if it is close to (away from) some control population.

(VI) Locally Optimal Procedures

For many classical procedures for selecting a subset containing the best population, the infimum of P(CS) occurs at the equiparameter points $\theta_1 = \ldots = \theta_k$. This provided the motivation for deriving rules which are optimal in some suitable sense in a neighborhood of an equiparameter configuration.

Let the independent observations $X_{i1},\ldots,X_{in_i}$ from π_i have pdf $f(.;\theta_i)$, $i = 1,\ldots,k$, where θ_i belongs to an open interval containing origin. A rank configuration is defined to be an N-tuple $\Delta = (\Delta_1,\ldots,\Delta_N)$, $\Delta_i \in \{1,\ldots,k\}$, where $N = \sum_{i=1}^{k} n_i$ and $\Delta_i = j$ means that the i-th smallest observation in the pooled sample comes from π_j. Define

$$B_j = \int_{-\infty}^{\infty} \int_{-\infty}^{x_N} \cdots \int_{-\infty}^{x_2} \left. \frac{\partial f(x_j;\theta)}{\partial \theta} \right|_{\theta=0} \prod_{\substack{i=1 \\ i\neq j}}^{N} f(x_i;0) \quad dx_1 \ldots dx_N \, .$$

and

$$A_i(\Delta) = \sum_{j:\Delta_j = i} B_j \, , \qquad V = \sum_{i=1}^{k} A_i(\Delta).$$

Let $P_{\underset{\sim}{\theta}}(\Delta)$ denote the probability of observing rank configuration Δ when the value of the parameter is $\underset{\sim}{\theta}$ with $\theta_i \neq 0$, $i = 1,\ldots,k$. A decision rule δ for subset selection will be called locally optimal if $\inf_{\underset{\sim}{\theta}\in\Omega_o} P_{\underset{\sim}{\theta}}(CS|\delta) = P^*$, a specified constant, and $P_{\underset{\sim}{\theta}}(CS|\delta)$ is maximum in a neighborhood ω_o of $\underset{\sim}{\theta}_o \in \Omega_o = \{ \underset{\sim}{\theta} : \theta_1 = \ldots = \theta_k \}$

Gupta, Huang and Nagel (1979) consider the problem of finding a permutation invariant locally optimal decision rule δ based on rank configuration Δ. Under certain regularity conditions on $f(.;\theta_i)$, $i = 1,\ldots,k$, it is shown that the decision rule δ for which

$$\psi_i^{\delta}(\Delta) = \begin{cases} 1 & , \text{ if } A_i(\Delta) > c \\ \rho_i & , \text{ if } A_i(\Delta) = c \\ 0 & , \text{ if } A_i(\Delta) < c \end{cases}$$

is a locally optimal, where constants ρ_i and c satisfy

$$\sum_{\Delta:A_i(\Delta)>c} P_{\underset{\sim}{\theta}_o}(\Delta) + \rho_i \sum_{\Delta:A_i(\Delta)=c} P_{\underset{\sim}{\theta}_o}(\Delta) = P^* .$$

It is also shown that, under the location parameter logistic model, the decision rule

$$R : \text{Select } \pi_i \text{ iff } H_i \geq d,$$

proposed by Gupta and McDonald (1970) is locally optimal, where H_i = average rank of observations from population π_i in the pooled sample, $i = 1,\ldots,k$.

Huang and Panchpakesan (1982) consider two types of goal namely, (i) select a subset of $\{\pi_1,\ldots,\pi_k\}$ containing UEP and (ii) select from $\pi_1,\ldots,\pi_k$, those populations, if any, for which the associated parameter values are larger than θ_o, the parameter of some control population. Consider the decision rule δ^* for which

$$\psi_i^{\delta^*}(\Delta) = \begin{cases} 1 & , \text{ if } A_i(\Delta) > V + D \\ \rho & , \text{ if } A_i(\Delta) = V + D \\ 0 & , \text{ if } A_i(\Delta) < V + D, \end{cases}$$

where D and ρ are chosen such that V+D > 0 and for $\underset{\sim}{\theta}_o \in \Omega_o$,

$$P_{\underset{\sim}{\theta}_o}(A_i(\Delta) > V+D) + \rho\, P_{\underset{\sim}{\theta}_o}(A_i(\Delta) = V+D) = P^*.$$

For the problem of selecting a subset containing UEP it is proved that decision rule δ^* is locally optimal. For logistic density δ^* reduces to the randomized version of the rule proposed by Gupta and McDonald (1970). For $i=1,\ldots,k$, let

$$A_i(\Delta,\underset{\sim}{\theta}) = \sum_{j:\Delta_j=i} \int_{-\infty}^{\infty} \int_{-\infty}^{x_N} \cdots \int_{-\infty}^{x_2} \frac{f(x_j;\theta_i) - f(x_j;0)}{\theta_i} \times \prod_{l=1}^{j-1} f(x_l;0) \prod_{l=j+1}^{N} f(x_l;\theta_{\Delta_l})\, dx_1 \ldots dx_N$$

For selecting populations better than control, it is assumed that θ_0 is unknown and there exists a known quantity θ_0^* such that $\theta_0 \le \theta_0^*$. For this problem, the decision rule δ^0 defined by

$$\psi_i^{\delta^0}(\Delta) = \begin{cases} 1 & , \text{ if } A_i(\Delta,\underset{\sim}{\theta}_0^*) > \frac{c_i}{\lfloor N} \\ \rho & , \text{ if } A_i(\Delta,\underset{\sim}{\theta}_0^*) = \frac{c_i}{\lfloor N} \\ 0 & , \text{ if } A_i(\Delta,\underset{\sim}{\theta}_0^*) < \frac{c_i}{\lfloor N} \end{cases}$$

where $\underset{\sim}{\theta}_0^* = (\theta_0^*,\ldots\theta_0^*)$ and $0 < \rho < 1$ and c_i are determined so that $\frac{1}{\lfloor N} \sum_{\Delta} \psi_i^{\delta^0}(\Delta) = P^*$, maximizes

$$\sum_{i=1}^{k} \frac{\partial}{\partial\theta_i} P_{\underset{\sim}{\theta}}(\pi_i \text{ is selected } | \theta_j = \theta_0^* < \theta_i,\ j \ne i)\Big|_{\theta_i=\theta_0^*}$$

among all procedures δ for which

$$P_{\underset{\sim}{\theta}}(\text{select } \pi_i | \underset{\sim}{\theta} \in \Omega_0, \delta) \le \gamma \qquad \text{for } i = 1,\ldots,k.$$

For selecting populations better than control, Huang, Panchpakesan and Tseng (1984) consider only those decision rules

for which $P_{\underset{\sim}{\theta}}(\pi_i$ is selected) $= \gamma_i$, $i = 1,\ldots,k$ for $\theta_0=\theta_1=\ldots=\theta_k$.

It is assumed that decision rule depends upon the observation through the statistics T_{io}, $i = 1,\ldots,k$, where T_{io} is suitably defined to indicate the difference between π_i and π_o. Let ξ_{io} be the parameter associated with pdf of T_{io}. The parameter ξ_{io} is a measure of separation between π_i and π_o. The local optimality criterion used amounts to maximizing

$$\sum_{i=1}^{k} \frac{\partial}{\partial \xi_{io}} E(\psi_i^{\delta}(\underset{\sim}{X}))\Big|_{(\xi_{1o}=\ldots=\xi_{ko}=\xi^*)}$$

where ξ^* is a specified constant.

(VII) Empirical Bayes Procedures

The first contribution to ranking and selection problems using empirical Bayes approach was made by Deely (1965). He assumed that $\tau(\underset{\sim}{\theta}) = \prod_{i=1}^{k} \tau_i(\theta_i)$ and derived empirical Bayes rules for the problem of selecting UEP. His loss function is given by

$$L(\underset{\sim}{\theta}, j) = c(\theta_{[k]} - \theta_j) \qquad (1.3.3)$$

where $c > 0$ is a constant.

Later Van Ryzin (1970) generalized the results of Deely (1965) to the case where $\Theta_{1j},\ldots\Theta_{kj}$ may not be independently distributed. For the loss function defined by (1.3.3), two empirical Bayes rules (depending on whether underlying pdf's are discrete or continuous) are proposed and are shown to be a.o. of order $1/\sqrt{n}$.

Gupta and Liang (1986a,b) derived empirical Bayes rules for selecting best binomial population and most probable multinomial

cell respectively, with respect to the loss function (1.3.3).

Some other contributors to this area are Gupta and Hsiao (1983) and Gupta and Leu (1983) who derived empirical Bayes rules for the problem of selecting good uniform populations. Also Gupta and Liang (1984) considered the problem of selecting good binomial populations and obtained empirical Bayes procedures.

(VIII) Asymptotic Consistent Procedures

Let δ_n be a selection rule based on a sample of size n with risk function $R_n(\underset{\sim}{\theta},\delta_n)$. Let $M(\underset{\sim}{\theta}) = \min\limits_{a \in \mathscr{A}} L(\underset{\sim}{\theta},a)$. Then the sequence of selection rules $\{\delta_n\}$ is said to be consistent at $\underset{\sim}{\theta}$ if $R_n(\underset{\sim}{\theta},\delta_n) \longrightarrow M(\underset{\sim}{\theta})$ as $n \to \infty$. We say that δ_n is point-wise consistent if δ_n is consistent at each $\underset{\sim}{\theta} \in \Omega$, while δ_n is called uniformly consistent if

$$\sup_{\underset{\sim}{\theta} \in K} [R_n(\underset{\sim}{\theta},\delta_n) - M(\underset{\sim}{\theta})] \longrightarrow 0 \text{ for all compact subsets } K \text{ of } \Omega.$$

Bjornstad (1984) obtained some necessary and sufficient conditions for both point-wise consistency and uniform consistency of permutation invariant selection rules. For the case when π_i is a normal population with mean θ_i and variance σ^2, following six loss functions were considered

$$\text{(i) } L_1\text{: } L_1^{\sigma}(\underset{\sim}{\theta},a) = \theta_{[k]} - \frac{1}{|a|}\sum_{j\in a}\theta_j + c(\theta_{[k]} - \max_{i\in a}\theta_i)$$

$$\text{(ii) } L_2\text{: } L_2^{\sigma}(\underset{\sim}{\theta},a) = \theta_{[k]} - \frac{1}{|a|}\sum_{j\in a}\theta_j + c\, I_{(\max_{i\in a}\theta_i < \theta_{[k]})}$$

$$\text{(iii) } L_3\text{: } L_3^{\sigma}(\underset{\sim}{\theta},a) = |a| + \theta_{[k]} - \max_{i\in a}\theta_i$$

$$\text{(iv) } L_4\text{: } L_4^{\sigma}(\underset{\sim}{\theta},a) = c_1\, I_{(\max_{i\in a}\theta_i < \theta_{[k]})} + c_2\,|a|$$

$$(v)\ L_5:\ L_5^{\sigma}(\underset{\sim}{\theta}, a) = |a| + c \sum_{i \notin a} I_{(\theta_i = \theta_{[k]})}$$

$$(vi)\ L_6:\ L_6^{\sigma}(\underset{\sim}{\theta}, a) = \sum_{i \in a} (\theta_{[k]} - \theta_i) + c \sum_{i \notin a} I_{(\theta_i = \theta_{[k]})}$$

where c, c_1, c_2 are positive constants. Let δ_n^* be a decision rule with associated individual selection probabilities

$$\psi_i^{\delta_n^*} = 1 \quad , \text{ if } \overline{X}_i^n \geq \overline{X}_{[k]}^n - S_n d_n$$
$$= 0 \quad , \text{ otherwise.} \qquad (1.3.4)$$

where $\overline{X}_i^n$ is the sample mean based on n observations from π_i, S_n^2 is the pooled sampled variance and d_n is a non-negative constant. Following results are proved

(i) Under the loss function L_1 (L_2), δ_n^* is uniformly (point-wise) consistent if and only if $d_n \longrightarrow 0$.

(ii) Under the loss function L_3 (L_4), δ_n^* is uniformly (point-wise) consistent if and only if $\sqrt{n}\, d_n \longrightarrow 0$.

(iii) Under the loss function L_5 with $c > 1$ (L_6) δ_n^* is point-wise consistent if and only if $d_n \longrightarrow 0$ ($\sqrt{n}\, d_n \longrightarrow 0$).

(iv) If $\inf_{\Omega \times A} P_{\underset{\sim}{\theta}, \sigma^2}(CS|\delta_n^*| = P^*$, $(\frac{1}{k} < P^* < 1)$ then δ_n^* is uniformly (point-wise) consistent under the loss function L_1 (L_2) but not consistent under any of the other loss functions given above.

Bjornstad also studied the consistency of admissible procedures derived by Bjornstad (1981a) and Bayes procedures derived by Chernoff and Yahav (1977), Goel and Rubin (1977) and Hsu (1978) under the respective loss functions considered by them.

Some necessary and sufficient conditions for point-wise and uniform consistency were obtained.

Bjornstad (1986) considered loss functions of the form

$$L_7: L_7(\underset{\sim}{\theta}, a) = \beta(|a|) \sum_{i \in a} \ell(\theta_{[k]} - \theta_i),$$

$$L_8: L_8(\underset{\sim}{\theta}, a) = \beta(|a|) \sum_{i \in a} \ell\left(\frac{\theta_{[k]}}{\theta_i} - 1\right),$$

and

$$L_9: L_9(\underset{\sim}{\theta}, a) = |a|,$$

where $\beta(i) > 0$, $i = 1, \ldots, k$ and ℓ is a continuous, non-decreasing function with $\ell(0) = 0$, and $\ell(x) > x$ if $x > 0$. Some necessary and sufficient conditions for point-wise and uniform consistency are developed for the class of decision rules satisfying P^*-condition. Applications are made to normal, multinomial and multivariate normal populations. It is shown that

(i) For selecting normal means δ_n^* defined by (1.3.4) is uniformly consistent. In fact Gupta's (1965) procedure is the only procedure in Seal's (1955) class which is uniformly consistent. The other classes of uniformly consistent procedures are the exponential procedures studied by Bjornstad (1981,a,b) and the class of procedures considered by Gupta and Panchpakesan (1972).

(ii) The rule proposed by Gupta and Sobel (1960) is uniformly consistent for selecting binomial populations while the conditional rule of Gupta and Nagel (1971) is not point-wise consistent.

(iii) Class of minimax rules studied by Berger (1980) for selecting multinomial cells is not consistent. However, an admissible and minimax rule proposed by Berger (1982) and the rules suggested by Gupta and Nagel (1967) and Gupta and Huang (1975) are uniformly consistent.

(iv) For selecting multivariate normal populations according to Mahalanobis distance the procedure investigated by Gupta (1966) is uniformly consistent, while the procedure based on maximum likelihood estimates considered by Alam and Rizvi (1966) is not.

Bjornstad (1985) studied asymptotic consistency of procedures for the problem of selecting good populations and obtained results similar to those of Bjornstad (1984, 1986).

(IX) Optimal Procedures Under Heteroscedasticity

Suppose that the observation X_i from π_i is distributed as normal with mean θ_i and variance σ_i^2, $i = 1,\ldots,k$. For the goal of selecting UEP, it has been established that if the risk is measured by probability of incorrect selection, the natural decision rule δ^N which selects population corresponding to the largest observation is minimax provided $\sigma_1^2 = \ldots = \sigma_k^2$. The most general version of this "Bahadur-Goodman-Lehmann Theorem" is presented in Gupta and Miescke (1984b). Suppose the assumption of equality of variances is dropped, then the situation changes drastically and no alternative of natural decision rule δ^N has been found so far which can be considered to be better in some reasonable sense.

When the risk is measured by probability of incorrect selection, Gupta and Miescke (1986) establish that natural decision rule is minimax if and only if $\sigma_1^2 = \ldots = \sigma_k^2$. Let

$$R(\underset{\sim}{\theta},\delta) = 1-P_{\underset{\sim}{\theta}}(CS|\delta) \tag{1.3.5}$$

$$\tilde{R}(\underset{\sim}{\theta},\delta) = \frac{1}{\underline{|k}} \sum_{g\in G} R(g\underset{\sim}{\theta},\delta) \tag{1.3.6}$$

and let $h_1,\ldots,h_k$ be k strictly increasing functions defined on real line. It is shown that if risk is measured by (1.3.6) then the decision rule δ^h which selects in terms of the largest $h_i(X_i)$, $i = 1,\ldots,k$ is minimax and the minimax value of the problem is $1 - \frac{1}{k}$. When the risk is given by (1.3.5) no alternative decision rule is proposed and Bayes rules with respect to various priors are studied to find reasonable modification or alternative to δ^N.

Now suppose that the observation X_i from π_i has a binomial distribution with single trial success probability θ_i and sample size n_i, $i = 1\ldots,k$. For the indifference-zone approach, Hall (1959) proved that the decision rule proposed by Sobel and Huyett (1957) is minimax if $n_1 = \ldots = n_k$ and if the loss function is $-\frac{1}{P^*}$ if a correct decision is made and zero otherwise, where P^* $(\frac{1}{k} < P^* < 1)$ is a pre-assigned constant. The decision rule proposed by Sobel and Huyett (1957) is also found to be most economical. It is pointed out that if assumption of equal sample sizes is dropped then problem becomes exceedingly complex.

For the case of unequal sample sizes, Gupta and Sobel (1958, 1960) have suggested choosing the population which yields the largest proportion of successes (randomizing in the case of ties)

as the population with the largest success probability. This will be referred to as the intuitive decision rule.

Risko (1985) considered the problem of finding a minimax rule for selecting the better of two binomial populations (the one with the larger single trial success probability) using indifference-zone approach when the risk is measured by probability of incorrect selection and the sample sizes are unequal. Let the observation X_i from π_i have binomial distribution with sample size n_i and single trial success probability θ_i, $i=1,2$. A decision rule δ^* is desired which maximizes the minimum P(CS) over the subset $\Omega(d) = \{ \underset{\sim}{\theta} : \theta_{[2]} \geq \theta_{[1]} + d \}$ of the parameter space, where $0 < d < 1$ is a pre-assigned number. It is observed thet the natural decision rule which selects the population corresponding to $\max(\frac{X_1}{n_1}, \frac{X_2}{n_2})$ as the better population performs very badly when one sample size is relatively larger than the other. When one sample size goes to infinity while the other is kept fixed, a minimax rule is derived. For the case when both the sample sizes n_1, n_2 are kept fixed, a class of decision rules depending on two parameters β_1 and β_2 ($2\beta_1 \leq 1+\beta_2$) is proposed and the minimax rules within this restricted class arre obtained numerically for some values of n_1 and n_2. The rules in the proposed class select the population π_1 as better population with probability $\psi_{\beta_1,\beta_2}(\frac{X_1}{n_1},\frac{X_2}{n_2}) = \varphi(\beta_1+(1-2\beta_1+\beta_2)\frac{X_1}{n_1}-\beta_2\frac{X_2}{2n_2})$, where

$$\varphi(x) = \begin{cases} 1, & \text{if } x > 1 \\ x, & \text{if } 0 \leq x \leq 1 \\ 0, & \text{if } x < 0. \end{cases}$$

It is shown that the restricted minimax rule in the proposed class is globally minimax if one of the sample sizes tends to infinity. It is also shown that the restricted minimax rule possesses many properties consistent with the globally minimax rule and is, in fact, globally minimax for certain combinations of sample sizes. For this problem of selecting the better of two binomial populations we give some necessary conditions for a rule to be minimax in Chapter V of this dissertation.

Recently Ambughalous and Miescke (1989) proved that the natural decision rule for selecting UEP is minimax if and only if $n_1 = \dots = n_k$, a result similar to Gupta and Miescke (1986). They also study the properties of Bayes rules under linear and general monotone permutation invariant loss.

(X) Procedures for Simultaneous Selection of Extreme Populations

Mishra and Dudewicz (1987) dealt with the problem of simultaneously selecting two subsets S_L and S_U such that S_L contains the LEP and S_U contains the UEP with a pre-assigned probability not less than P^* ($\frac{1}{k(k-1)} < P^* < 1$). Let the observations from π_i have density $f(x-\theta_i)$, $i = 1,\dots,k$ and suppose that f has the MLR property. For the above goal, they propose the following procedure R_{MD}.

R_{MD} : Select π_i in S_L iff $X_i \le X_{[1]} + d_1$, and

Select π_i in S_U iff $X_i \ge X_{[k]} - d_2$,

where d_1 and d_2 are positive constants such that

$$\inf_{\Omega} P_{\underset{\sim}{\theta}}(CS|R_{MD}) = \inf_{\Omega} P_{\underset{\sim}{\theta}}(\pi_{(1)} \in S_L, \pi_{(k)} \in S_U|R_{MD}) = P^*.$$

Note that Gupta's (1965) procedure R_2 becomes a special case of

R_{MD} by taking $d_1 = \infty$. It is proved that

(i) $\inf_{\Omega} P_{\underset{\sim}{\theta}}(CS|R_{MD})$ and $\sup_{\Omega} E_{\underset{\sim}{\theta}}(|S_L| + |S_U| \big| R_{MD})$ occur at $\theta_1 = \ldots = \theta_k$ and $2kP^* \le E_{\underset{\sim}{\theta}}(|S_L| + |S_U| \big| R_{MD}) \le 2k$.

(ii) If $\theta_i \ge \theta_j$, then $P_{\underset{\sim}{\theta}}(\pi_j \in S_L|R_{MD}) \ge P_{\underset{\sim}{\theta}}(\pi_i \in S_L|R_{MD})$ and $P_{\underset{\sim}{\theta}}(\pi_i \in S_U|R_{MD}) \ge P_{\underset{\sim}{\theta}}(\pi_j \in S_U|R_{MD})$.

For $d_1 = d_2 = d$(say), the lower and upper bounds on the infimum of $P_{\underset{\sim}{\theta}}(CS|R_{MD})$ are obtained to approximate d from the existing tables of Gupta (1965). Application to normal means problem with known and equal variances is made. For simultaneous selection of subsets containing LEP and UEP in the case of normal populations with means $\theta_1, \ldots, \theta_k$ and unknown variances $\sigma_1^2, \ldots \sigma_k^2$, a two-stage procedure similar to that of Dudewicz and Dalal (1975) is also proposed. Tables for applications are provided.

<u>Mishra</u> (1986a,b,c) considered some variations of this problem under different population models, viz, (i) simultaneous selection of subsets containing LEP and UEP with variances of normal distributions as the parameters of interest, (ii) simultaneous selection of LEP and UEP using indifference zone approach with means of normal distributions as the parameters of interest, (iii) non-parametric simultaneous selection of subsets containing LEP and UEP when LEP (UEP) is the population with the smallest (largest) α-quantile.

CHAPTER II

SIMULTANEOUS SELECTION OF EXTREME POPULATIONS : A BAYESIAN APPROACH

2.1 Introduction

Let $X_1,\dots,X_k$ denote $k(\geq 2)$ independent random variables representing the populations $\pi_1,\dots,\pi_k$, respectively, and suppose that X_i has a pdf $f(.;\theta_i)$, $i = 1,\dots,k$. Recall that the populations $\pi_{(1)}$ and $\pi_{(k)}$ associated with $\theta_{[1]} = \min(\theta_1,\dots,\theta_k)$ and $\theta_{[k]} = \max(\theta_1,\dots,\theta_k)$, respectively, are called the lower extreme population (LEP) and the upper extreme population (UEP). Mishra and Dudewicz (1987) considered the problem of simultaneously selecting two non-empty subsets S_L and S_U containing LEP and UEP, respectively, and gave a procedure under the assumption that θ_i is a location parameter and $f(.;\theta_i)$ has the MLR property, $i = 1,\dots,k$. In this chapter we approach this problem from a Bayesian point of view extending the results of Goel and Rubin (1977) who studied the problem of selecting a non-empty subset containing UEP for a specific loss function.

Let $\tau(\underset{\sim}{\theta})$ be a prior distribution on the parameter space Ω and suppose that $\mathcal{D}$ is the class of all decision rules. For the problem at hand the action space is

$$\mathcal{A} = \{(a_1,a_2) : a_1 \subset \{1,\dots,k\},\ a_2 \subset \{1,\dots,k\},\ a_1 \neq \phi,\ a_2 \neq \phi\}$$

where ϕ denotes the empty set. If $a_1 = \{i_1,\dots,i_l\}$ and

$a_2 = \{j_1,\ldots,j_m\}$, then taking an action (a_1,a_2) means selecting $\{\pi_{i_1},\ldots,\pi_{i_\ell}\}$ as S_L and $\{\pi_{j_1},\ldots,\pi_{j_m}\}$ as S_U.

Let $L(\underset{\sim}{\theta},(a_1,a_2))$ be the loss incurred in taking action (a_1,a_2) when $\underset{\sim}{\theta}$ is the true value of parameter. The problem in general terms then is to find a decision rule $\delta_\tau \in \mathcal{D}$ which minimizes the Bayes risk

$$r(\tau,\delta) = \int_\Omega \int_{\mathcal{X}} \sum_{\mathcal{A}} \delta((a_1,a_2)|\underset{\sim}{x})\, L(\underset{\sim}{\theta},(a_1,a_2)) \left[\prod_{i=1}^{k} f(x_i;\theta_i)\right] d\tau(\underset{\sim}{\theta})\, d\underset{\sim}{x}$$

for the given loss function L and prior distribution τ, where $d\underset{\sim}{x} = dx_1\ldots dx_n$. Such a rule δ_τ is called a Bayes decision rule with respect to the prior τ and loss function L.

In Sections 2.2 and 2.3 we consider the case when the selected subsets need not be disjoint whereas in Section 2.4 we impose the requirement that the selected subsets should be disjoint. In Section 2.5 we apply the results of Section 2.4 to the case where the underlying distributions are assumed to be normal. We determine an essentially complete class of decision rules for finding Bayes rules with respect to a general loss function in Section 2.2. In Section 2.3 the loss function is assumed to be semi-additive and non-negative and in Sections 2.4 and 2.5 a specific loss function is considered.

We assume throughout that the assumptions of Fubini's Theorem are satisfied so that for every $\delta \in \mathcal{D}$

$$r(\tau,\delta) = \int_{\mathcal{X}} \sum_{\mathcal{A}} \delta((a_1,a_2)|\underset{\sim}{x}) \left\{ \int_{\Omega} L(\underset{\sim}{\theta},(a_1,a_2)) \left[\prod_{i=1}^{k} f(x_i;\theta_i) \right] d\tau(\underset{\sim}{\theta}) \right\} d\underset{\sim}{x}$$

$$= \int_{\mathcal{X}} \sum_{\mathcal{A}} \delta((a_1,a_2)|\underset{\sim}{x})\; r_1(\underset{\sim}{x},(a_1,a_2))\; d\underset{\sim}{x}, \tag{2.1.1}$$

where

$$r_1(\underset{\sim}{x},(a_1,a_2)) = \int_{\Omega} L(\underset{\sim}{\theta},(a_1,a_2)) \left[\prod_{i=1}^{k} f(x_i;\theta_i) \right] d\tau(\underset{\sim}{\theta}) \tag{2.1.2}$$

2.2 An Essentially Complete Class For Bayes Rules

Let $(a_1^*,a_2^*) \in \mathcal{A}$ be such that

$$r_1(\underset{\sim}{x},(a_1^*,a_2^*)) = \min_{(a_1,a_2)\in\mathcal{A}} r_1(\underset{\sim}{x},(a_1,a_2)), \quad \forall\, \underset{\sim}{x} \in \mathcal{X}$$

where $r_1(\underset{\sim}{x},(a_2,a_2))$ is given by (2.1.2).

Lemma 2.2.1 : Let τ be a symmetric prior distribution on the parameter space Ω and suppose that the loss function L is permutation invariant, that is $L(\underset{\sim}{\theta},(a_1,a_2)) = L(g\underset{\sim}{\theta},(ga_1,ga_2))$, $\forall\, g \in G$, where G is the group of permutations on $\{1,\ldots,k\}$. Then δ_τ given by

$$\delta_\tau((ga_1^*,ga_2^*)|g\underset{\sim}{x}) = \begin{cases} \dfrac{1}{|G_{\underset{\sim}{x}}|}, & \forall\, g \in G_{\underset{\sim}{x}},\ \underset{\sim}{x} \in \mathcal{X} \\ 0, & \text{otherwise} \end{cases}$$

minimizes $r(\tau,\delta)$ over $\mathcal{D}$, where $G_{\underset{\sim}{x}} = \{g \in G : g\underset{\sim}{x} = \underset{\sim}{x}\}$, and $|G_{\underset{\sim}{x}}|$ denotes the cardinality of the set $G_{\underset{\sim}{x}}$.

Proof: Using invariance of L and symmetry of τ we have

$$r_1(\underset{\sim}{x},(ga_1^*,ga_2^*)) = \int_\Omega L(\underset{\sim}{\theta},(ga_1^*,ga_2^*)) \prod_{i=1}^k f(x_i;\theta_i)\, d\tau(\underset{\sim}{\theta})$$

$$= \int_\Omega L(g^{-1}\underset{\sim}{\theta},(a_1^*,a_2^*)) \prod_{i=1}^k f(x_{gi};\theta_{gi})\, d\tau(g^{-1}\underset{\sim}{\theta})$$

$$= \int_\Omega L(\underset{\sim}{\theta},(a_1^*,a_2^*)) \prod_{i=1}^k f(x_{gi};\theta_i)\, d\tau(\underset{\sim}{\theta})$$

$$= r_1(g^{-1}\underset{\sim}{x},\ (a_1^*,a_2^*))$$

$$= r_1(\underset{\sim}{x},(a_1^*,a_2^*)), \qquad \text{if } \underset{\sim}{x} = g\underset{\sim}{x}.$$

Hence, if $\underset{\sim}{x} = g\underset{\sim}{x}$

$$r_1(\underset{\sim}{x},(a_1^*,a_2^*)) = r_1(\underset{\sim}{x},(ga_1^*,ga_2^*)) \tag{2.2.1}$$

Now, let δ be any decision rule. Then

$$r(\tau,\delta) = \int_{\mathcal{X}} \sum_{\mathcal{A}} \delta((a_1,a_2)|\underset{\sim}{x})\ r_1(\underset{\sim}{x},(a_1,a_2))\ d\underset{\sim}{x}$$

$$\geq \int_{\mathcal{X}} r_1(\underset{\sim}{x},(a_1^*,a_2^*))d\underset{\sim}{x}$$

Also,

$$r(\tau,\delta_\tau) = \int_{\mathcal{X}} \sum_{\mathcal{A}} \delta_\tau((a_1,a_2)|\underset{\sim}{x})\ r_1(\underset{\sim}{x},(a_1,a_2))d\underset{\sim}{x}$$

$$= \int_{\mathcal{X}} \sum_{g:g\in G_{\underset{\sim}{x}}} \delta_\tau((ga_1^*,ga_2^*)|\underset{\sim}{x})\ r_1(\underset{\sim}{x},(ga_1^*,ga_2^*))d\underset{\sim}{x}$$

$$= \int_{\mathcal{X}} r_1(\underset{\sim}{x},(a_1^*,a_2^*))d\underset{\sim}{x} \qquad \text{(using (2.2.1))}.$$

Hence, $r(\tau,\delta) \geq r(\tau,\delta_\tau)$, $\forall\ \delta \in \mathcal{D}$. ■

Let A^c denote the complement of a set A. Define

$B(a_1,a_2) = \{\underset{\sim}{x} \in \mathcal{X} : x_i \leq x_l \leq x_j \text{ for } i \in a_1,\ j \in a_2,\ l \in a_1^c \cap a_2^c\}$

$$\mathscr{A}(s,t) = \{(a_1,a_2) : |a_1| = s,\ |a_2| = t,\ 1 \le s,t \le k\}, \text{ and}$$

$$H_{\underset{\sim}{x}}(s,t) = \{(a_1,a_2) : (a_1,a_2) \in \mathscr{A}(s,t),\ \underset{\sim}{x} \in B(a_1,a_2)\}.$$

Let $\{i_1,\ldots,i_s,l\}$ and $\{j_1,\ldots,j_t,m\}$ denote subsets of indices belonging to the set $\{1,\ldots,k\}$, $1 \le s,\ t \le k-1$ and assume that the loss function L satisfies the following conditions :

$$\left.\begin{array}{l}
\text{(i) } L(\underset{\sim}{\theta},(a_1,a_2)) \ge 0,\ \forall\ \underset{\sim}{\theta} \in \Omega,\ \forall (a_1,a_2) \in \mathscr{A}, \\
\text{(ii) if } \theta_j \le \theta_i, \text{ then} \\
\qquad L(\underset{\sim}{\theta},(a_1,\{j_1,\ldots,j_t,j\})) \ge L(\underset{\sim}{\theta},(a_1,\{j_1,\ldots,j_t,i\})) \\
\text{(iii) if } \theta_j \le \theta_i, \text{ then} \\
\qquad L(\underset{\sim}{\theta},(\{i_1,\ldots,i_s,i\},a_2)) \ge L(\underset{\sim}{\theta},(\{i_1,\ldots,i_s,j\},a_2)) \\
\text{(iv) } \forall\ \underset{\sim}{\theta} \in \Omega, \\
\qquad L(\underset{\sim}{\theta},(\{i_1,\ldots,i_s,i\},\{j_1,\ldots,j_t,i\})) + \\
\qquad L(\underset{\sim}{\theta},(\{i_1,\ldots,i_s,j\},\{j_1,\ldots,j_t,j\}) \\
\qquad = L(\underset{\sim}{\theta},(\{i_1,\ldots,i_s,i\},\{j_1,\ldots,j_t,j\})) + \\
\qquad L(\underset{\sim}{\theta},(\{i_1,\ldots,i_s,j\},\{j_1,\ldots,j_t,i\})
\end{array}\right\} \quad (2.2.2)$$

Lemma 2.2.2 : Under the assumptions of Lemma 2.2.1, if the densities $f(\cdot,;\theta_i)$ $(1 \le i \le k)$ have MLR property and the loss function L satisfies conditions (i) - (iv) of (2.2.2), then for any (a_1,a_2), $(a_1',a_2') \in \mathscr{A}(s,t)$ we have

$$r_1(\underset{\sim}{x},(a_1,a_2)) \le r_1(\underset{\sim}{x},(a_1',a_2')), \qquad \forall\ \underset{\sim}{x} \in B(a_1,a_2).$$

Proof: Let (a_1,a_2), $(a_1',a_2') \in \mathscr{A}(s,t)$ and suppose that $|a_2 \cap a_2'| = q$, that is, a_2 and a_2' have q elements in common. Without loss of generality assume that

$$a_2 = \{j_1,\dots,j_q,\ j_{q+1},\dots,j_t\}$$

and

$$a_2' = \{j_1,\dots,j_q,\ m_{q+1},\dots,m_t\}.$$

Let

$$a_1' = \{l_1,\dots,l_s\} \text{ and } a_2'' = \{j_1,\dots,j_q, j_{q+1},\ m_{q+2},\dots,m_t\}.$$

We first prove that

$$r_1(\underset{\sim}{x},(a_1',a_2'')) \le r_1(\underset{\sim}{x},(a_1',a_2')),\ \forall\ \underset{\sim}{x} \in B(a_1,a_2)$$

by considering the four possible cases.

<u>Case I</u> : $m_{q+1},\ j_{q+1} \in a_1'$

Without loss of generality assume that $l_{s-1} = m_{q+1},\ l_s = j_{q+1}$. Then

$$r_1(\underset{\sim}{x},(a_1',a_2')) - r_1(\underset{\sim}{x},(a_1',a_2''))$$

$$= \int_{\{\underset{\sim}{\theta}:\theta_{j_{q+1}} > \theta_{m_{q+1}}\}} [L(\underset{\sim}{\theta},(a_1',a_2'))-L(\underset{\sim}{\theta},(a_1',a_2''))] \prod_{i=1}^{k} f(x_i;\theta_i)d\tau(\underset{\sim}{\theta})$$

$$+ \int_{\{\underset{\sim}{\theta}:\theta_{j_{q+1}} = \theta_{m_{q+1}}\}} [L(\underset{\sim}{\theta},(a_1',a_2'))-L(\underset{\sim}{\theta},(a_1',a_2''))] \prod_{i=1}^{k} f(x_i;\theta_i)d\tau(\underset{\sim}{\theta})$$

$$+ \int_{\{\underset{\sim}{\theta}:\theta_{j_{q+1}} < \theta_{m_{q+1}}\}} [L(\underset{\sim}{\theta},(a_1',a_2'))-L(\underset{\sim}{\theta},(a_1',a_2''))] \prod_{i=1}^{k} f(x_i;\theta_i)d\tau(\underset{\sim}{\theta})$$

Now since L is permutation invariant and τ is symmetric, the second integral is zero and roles of $\theta_{j_{q+1}}$ and $\theta_{m_{q+1}}$ can be interchanged in the third integral so that

$$r_1(\underset{\sim}{x},(a_1',a_2')) - r_1(\underset{\sim}{x},(a_1',a_2''))$$

$$= \int_{\{\underset{\sim}{\theta}:\theta_{j_{q+1}}>\theta_{m_{q+1}}\}} [L(\underset{\sim}{\theta},(a_1',a_2'))-L(\underset{\sim}{\theta},(a_1',a_2''))] \prod_{\substack{i=1\\ i\neq j_{q+1},m_{q+1}}}^{k} f(x_i;\theta_i)\,[f(x_{j_{q+1}};\theta_{j_{q+1}})f(x_{m_{q+1}};\theta_{m_{q+1}})$$

$$- f(x_{j_{q+1}};\theta_{m_{q+1}})f(x_{m_{q+1}};\theta_{j_{q+1}})]\,d\tau(\underset{\sim}{\theta}).$$

Since $j_{q+1} \in a_2$, $m_{q+1} \notin a_2$ and $\underset{\sim}{x} \in B(a_1,a_2)$, it follows that $x_{j_{q+1}} \geq x_{m_{q+1}}$. Now using MLR property and (2.2.2(ii)), we get

$$r_1(\underset{\sim}{x},(a_1',a_2')) - r_1(\underset{\sim}{x},(a_1',a_2'')) \geq 0 \qquad \forall\ \underset{\sim}{x} \in B(a_1,a_2)).$$

<u>**Case II**</u> : $m_{q+1} \in a_1'$, $j_{q+1} \notin a_1'$.

Without loss of generality, assume that $l_s = m_{q+1}$. Proceeding as in Case I and using (2.2.2 (iv)) in addition to the MLR property and (2.2.2 (ii)), we get

$$r_1(\underset{\sim}{x},(a_1',a_2')) - r_1(\underset{\sim}{x},(a_1',a_2''))$$

$$= \int_{\{\underset{\sim}{\theta}:\theta_{j_{q+1}}>\theta_{m_{q+1}}\}} [L(\underset{\sim}{\theta},(a_1',a_2'))-L(\underset{\sim}{\theta},(a_1',a_2''))] \prod_{\substack{i=1\\ i\neq j_{q+1},m_{q+1}}}^{k} f(x_i;\theta_i)\,[f(x_{j_{q+1}};\theta_{j_{q+1}})\ f(x_{m_{q+1}};\theta_{m_{q+1}})$$

$$- f(x_{j_{q+1}};\theta_{m_{q+1}})\ f(x_{m_{q+1}};\theta_{j_{q+1}})]\ d\tau(\underset{\sim}{\theta})$$

$$\geq 0, \quad \forall\ \underset{\sim}{x} \in B(a_1,a_2).$$

<u>**Case III**</u> : $j_{q+1} \in a_1'$, $m_{q+1} \notin a_1'$.

Assuming $l_s = j_{q+1}$, without loss of generality, and proceeding on the lines of Case II we get

$$r_1(\underset{\sim}{x},(a'_1,a'_2)) - r_1(\underset{\sim}{x},(a'_1,a''_2))$$

$$= \int_{\{\underset{\sim}{\theta}:\theta_{j_{q+1}}>\theta_{m_{q+1}}\}} [L(\underset{\sim}{\theta},(a'_1,a'_2))-L(\underset{\sim}{\theta},(a'_1,a''_2))]$$

$$\prod_{\substack{i=1\\ i\neq j_{q+1},m_{q+1}}}^{k} f(x_i;\theta_i)\ [f(x_{j_{q+1}};\theta_{j_{q+1}})\ f(x_{m_{q+1}};\theta_{m_{q+1}})$$

$$- f(x_{j_{q+1}};\theta_{m_{q+1}})\ f(x_{m_{q+1}};\theta_{j_{q+1}})]\ d\tau(\underset{\sim}{\theta})$$

$$\geq 0, \quad \forall\ \underset{\sim}{x} \in B(a_1,a_2).$$

<u>**Case IV**</u> **:** $j_{q+1},\ m_{q+1} \notin a'_1$.

Again, using the invariance of L and symmetry of τ, we get

$$r_1(\underset{\sim}{x},(a'_1,a'_2)) - r_1(\underset{\sim}{x},(a'_1,a''_2))$$

$$= \int_{\{\underset{\sim}{\theta}:\theta_{j_{q+1}}>\theta_{m_{q+1}}\}} [L(\underset{\sim}{\theta},(a'_1,a'_2))-L(\underset{\sim}{\theta},(a'_1,a''_2))]$$

$$\prod_{\substack{i=1\\ i\neq j_{q+1},m_{q+1}}}^{k} f(x_i;\theta_i)\ [f(x_{j_{q+1}};\theta_{j_{q+1}})\ f(x_{m_{q+1}};\theta_{m_{q+1}})$$

$$- f(x_{j_{q+1}};\theta_{m_{q+1}})\ f(x_{m_{q+1}};\theta_{j_{q+1}})]\ d\tau(\underset{\sim}{\theta})$$

$$\geq 0, \quad \forall\ \underset{\sim}{x} \in B(a_1,a_2).$$

Combining Cases I - IV, we have

$$r_1(\underset{\sim}{x},(a'_1,a'_2)) \geq r_1(\underset{\sim}{x},(a'_1,a''_2)) \qquad \forall\ \underset{\sim}{x} \in B(a_1,a_2).$$

Proceeding in similar fashion, we get

$$r_1(\underset{\sim}{x},(a_1',a_2')) = r_1(\underset{\sim}{x},(\{l_1,\dots,l_s\},\{j_1,\dots,j_q,m_{q+1},\dots,m_t\}))$$
$$\geq r_1(\underset{\sim}{x},(\{l_1,\dots,l_s\},\{j_1,\dots,j_{q+1},m_{q+2},\dots,m_t\}))$$
$$\dots \quad \dots \quad \dots \quad \dots \quad \dots$$
$$\geq r_1(\underset{\sim}{x},(\{l_1,\dots,l_s\},\{j_1,\dots,j_t\}))$$
$$= r_1(\underset{\sim}{x},(a_1',a_2)), \quad \forall\ \underset{\sim}{x} \in B(a_1,a_2) \qquad (2.2.3)$$

Now keeping a_2 fixed and repeating the above arguments for indices in a_1 and using (2.2.2 (iii) - (iv)) we get

$$r_1(\underset{\sim}{x},(a_1',a_2)) \geq r_1(\underset{\sim}{x},(a_1,a_2)), \quad \forall\ \underset{\sim}{x} \in B(a_1,a_2) \qquad (2.2.4)$$

On combining (2.2.3) and (2.2.4), we get the result. ■

Corollary 2.2.1: Suppose that the assumptions of Lemma 2.2.2 are satisfied. Then $(a_1(\underset{\sim}{x}),\ a_2(\underset{\sim}{x})) \in H_{\underset{\sim}{x}}(s,t)$ implies that

$$r_1(\underset{\sim}{x},(a_1(\underset{\sim}{x}),a_2(\underset{\sim}{x}))) \leq r_1(\underset{\sim}{x},(a_1',a_2'))$$

for every $(a_1',a_2') \in \mathscr{A}(s,t)$.

Proof : Since $(a_1(\underset{\sim}{x}),\ a_2(\underset{\sim}{x})) \in H_{\underset{\sim}{x}}(s,t)$, implies that $\underset{\sim}{x} \in B(a_1(\underset{\sim}{x}),\ a_2(\underset{\sim}{x}))$, we have, by Lemma 2.2.2,

$$r_1(\underset{\sim}{x},(a_1(\underset{\sim}{x}),a_2(\underset{\sim}{x})) \leq r_1(\underset{\sim}{x},(a_1',a_2')), \qquad \forall (a_1',a_2') \in \mathscr{A}(s,t) \quad ■$$

For $s,t = 1,\dots,k$, let

$$a_{\{s\}} = \{\{1\},\dots,\{s\}\},\ a^{\{t\}} = \{\{k-t+1\},\dots,\{k\}\},$$

that is, $a_{\{s\}}$ and $a^{\{t\}}$ are subsets of indices of populations which correspond to s smallest and t largest observations, respectively. Define decision rule $\delta^{s,t}$ $(1 \leq s,t \leq k)$ as follows

$$\delta^{s,t}((a_1,a_2)|\underset{\sim}{x}) = \begin{cases} 1 & , \text{ if } (a_1,a_2) = (a_{\{s\}},a^{\{t\}}) \\ 0 & , \text{ otherwise} \end{cases}$$

Let

$$D_B = \{\delta^{s,t}, \quad s = 1,\dots,k,\ t = 1,\dots,k\}$$

and

$$\mathcal{D}_B = \{\delta \in \mathcal{D} : \sum_{s=1}^{k} \sum_{t=1}^{k} \delta((a_{\{s\}}, a^{\{t\}}) | \underset{\sim}{x}) = 1\}.$$

Theorem 2.2.1 : Under the assumptions of Lemma 2.2.2

(i) Class D_B is essentially complete for finding Bayes rule in the class D of all non-randomized rules.

(ii) Class $\mathcal{D}_B$ is essentially complete for finding Bayes rule in the class $\mathcal{D}$ of all randomized rules.

Proof : (i) From (2.1.1) it is clear that non-randomized Bayes rule selects subset pair (a_1^*, a_2^*) w.p. 1,

where

$$r_1(\underset{\sim}{x}, (a_1^*, a_2^*)) = \min_{(a_1\ a_2) \in \mathcal{A}} r_1(\underset{\sim}{x}, (a_1, a_2)), \quad \forall\ \underset{\sim}{x} \in \mathcal{X}$$

and from Corollary 2.2.1, we have

$$r_1(\underset{\sim}{x}, (a_{\{s\}}, a^{\{t\}})) \le r_1(\underset{\sim}{x}, (a_1', a_2')),$$

$$\text{for all } (a_1', a_2') \in \mathcal{A}(s,t),\ \underset{\sim}{x} \in \mathcal{X}$$

Hence the result follows.

(ii) Let δ be any decision rule. Consider the decision rule δ^* defined as follows : for $\underset{\sim}{x} \in \mathcal{X}$

$$\delta^*((a_1, a_2) | \underset{\sim}{x}) = \begin{cases} \displaystyle\sum_{(a_1', a_2') \in A(s,t)} \delta((a_1', a_2') | \underset{\sim}{x}), & \text{if } (a_1, a_2) = (a_{\{s\}}, a_{\{t\}}) \\ & s,t = 1,\dots,k \\ 0 & , \quad \text{otherwise} \end{cases}$$

Note that

$$\sum_{s=1}^{k} \sum_{t=1}^{k} \delta^*((a_{\{s\}},a^{\{t\}})|\underset{\sim}{x}) = \sum_{s=1}^{k} \sum_{t=1}^{k} \sum_{(a_1,a_2)\in\mathscr{A}(s,t)} \delta((a_1,a_2)|\underset{\sim}{x})$$

$$= \sum_{(a_1,a_2)\in\mathscr{A}} \delta((a_1,a_2)|\underset{\sim}{x})$$

$$= 1.$$

Hence $\delta^* \in \mathscr{D}_B$.

Also from Corollary 2.2.1,

$$r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{t\}})) \le r_1(\underset{\sim}{x},(a_1',a_2')),$$

$$\text{for all } (a_1',a_2') \in \mathscr{A}(s,t),\ \underset{\sim}{x} \in \mathscr{X}$$

Thus,

$$r(\tau,\delta) = \int_{\mathscr{X}} \sum_{(a_1,a_2)\in\mathscr{A}} \delta((a_1,a_2)|\underset{\sim}{x})\ r_1(\underset{\sim}{x},(a_1,a_2))\ d\underset{\sim}{x}$$

$$= \int_{\mathscr{X}} \sum_{s=1}^{k} \sum_{t=1}^{k} \left\{ \sum_{(a_1,a_2)\in\mathscr{A}(s,t)} \delta((a_1,a_2)|\underset{\sim}{x})\ r_1(\underset{\sim}{x},(a_1,a_2)) \right\} d\underset{\sim}{x}$$

$$\ge \int_{\mathscr{X}} \sum_{s=1}^{k} \sum_{t=1}^{k} \left\{ \sum_{(a_1,a_2)\in\mathscr{A}(s,t)} \delta((a_1,a_2)|\underset{\sim}{x})\ r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{t\}})) \right\} d\underset{\sim}{x}$$

$$= \int_{\mathscr{X}} \sum_{s=1}^{k} \sum_{t=1}^{k} \left\{ \sum_{(a_1,a_2)\in\mathscr{A}(s,t)} \delta^*(a_{\{s\}},a^{\{t\}})\ r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{t\}})) d\underset{\sim}{x} \right.$$

$$= \int_{\mathscr{X}} \left\{ \sum_{(a_1,a_2)\in\mathscr{A}} \delta^*((a_1,a_2)|\underset{\sim}{x})\ r_1(\underset{\sim}{x},(a_1,a_2)) \right\} d\underset{\sim}{x}$$

$$= r(\tau,\delta^*)$$

and the result follows. ■

2.3 Semi-additive and Non-negative Loss Function

For, $i, j = 1,\dots,k$, let $\ell_{i,j}(\underset{\sim}{\theta})$ be a non-negative function on the parameter space Ω satisfying the following conditions :-

$$\left.\begin{array}{ll} \text{(i)} & \ell_{gi,gj}(g\underset{\sim}{\theta}) = \ell_{i,j}(\underset{\sim}{\theta}), \qquad , \forall\, g \in G \\ \text{(ii)} & \theta_j \ge \theta_{j'} \implies \ell_{i,j}(\underset{\sim}{\theta}) \le \ell_{i,j'}(\underset{\sim}{\theta}) \\ \text{(iii)} & \theta_i \ge \theta_{i'} \implies \ell_{i',j}(\underset{\sim}{\theta}) \le \ell_{i,j}(\underset{\sim}{\theta}) \\ \text{(iv)} & \ell_{i,i}(\underset{\sim}{\theta}) + \ell_{j,j}(\underset{\sim}{\theta}) = \ell_{i,j}(\underset{\sim}{\theta}) + \ell_{j,i}(\underset{\sim}{\theta}) \end{array}\right\} \qquad (2.3.1)$$

Let

$$r_{i,j}(\underset{\sim}{x}) = \int_\Omega \ell_{i,j}(\underset{\sim}{\theta}) \prod_{l=1}^{k} f(x_l;\theta_l)\, d\tau(\underset{\sim}{\theta}).$$

Lemma 2.3.1 : Suppose that the pdf's $f(.;\theta_i)$, $i = 1,\dots,k$ possess MLR property, prior distribution τ is symmetric on Ω and $\ell_{i,j}$'s satisfy (i) - (iv) of (2.3.1). Then

$$r_{i,j}(\underset{\sim}{x}) \ge r_{\{1\},\{k\}}(\underset{\sim}{x}).$$

Proof : Let $\bar{L}$ be a permutation invariant loss function which satisfies conditions (2.2.3) and is such that for $i,j = 1,\dots,k$,

$$\bar{L}(\underset{\sim}{\theta},(\{i\},\{j\})) = \ell_{i,j}(\underset{\sim}{\theta}).$$

Therefore, on taking $s = t = 1$ in Corollary 2.2.1, we have for any $i,j = 1,\dots,k$

$$r_{i,j}(\underset{\sim}{x}) = \int_\Omega \ell_{i,j}(\underset{\sim}{\theta}) \prod_{l=1}^{k} f(x_l;\theta_l)\, d\tau(\underset{\sim}{\theta})$$

$$= \int_\Omega \bar{L}(\underset{\sim}{\theta},(\{i\},\{j\})) \prod_{l=1}^{k} f(x_l;\theta_l)\, d\tau(\underset{\sim}{\theta})$$

$$\geq \int_{\Omega} \bar{L}(\underset{\sim}{\theta},(\{\{1\}\},\{\{k\}\})) \prod_{l=1}^{k} f(x_l;\theta_l)\, d\tau(\underset{\sim}{\theta})$$

$$= \int_{\Omega} \ell_{\{1\},\{k\}}(\underset{\sim}{\theta}) \prod_{l=1}^{k} f(x_l;\theta_l)\, d\tau(\underset{\sim}{\theta})$$

$$= r_{\{1\},\{k\}}(\underset{\sim}{x}). \ \blacksquare$$

Let $\beta(i,j)$ be a non-negative function defined on $\mathbb{N} \times \mathbb{N}$, where $\mathbb{N}$ denotes the set of natural numbers. For the problem of simultaneously selecting two non-empty subsets S_L and S_U containing LEP and UEP, respectively, consider the loss function defined by

$$L(\underset{\sim}{\theta},(a_1,a_2)) = \beta(|a_1|,|a_2|) \sum_{i\in a_1}^{k} \sum_{j\in a_2}^{k} \ell_{i,j}(\underset{\sim}{\theta}) \tag{2.3.2}$$

we have

$$r_1(\underset{\sim}{x},(a_1,a_2)) = \int_{\Omega} \beta(|a_1|,\ |a_2|) \sum_{i\in a_1} \sum_{j\in a_2} \ell_{i,j}(\underset{\sim}{\theta}) \prod_{l=1}^{k} f(x_l;\theta_l) d\tau(\underset{\sim}{\theta})$$

$$= \beta(|a_1|,|a_2|) \sum_{i\in a_1} \sum_{j\in a_2} r_{i,j}(\underset{\sim}{x}).$$

__**Lemma 2.3.2**__ **:** Under the assumptions of Lemma 2.3.1,

$$r_1(\underset{\sim}{x},(\{\{1\}\},\ \{\{k\}\})) \leq r_1(\underset{\sim}{x},(a_1,a_2)), \quad \forall\ (a_1,a_2) \in \mathcal{A}$$

provided st $\beta(s,t) \geq \beta(1,1)$, $s,t = 1,\ldots,k$.

Proof : Suppose $(a_1, a_2) \in \mathscr{A}(s,t)$. Then

$$\begin{aligned} r_1(\underset{\sim}{x}, (a_1, a_2)) &= \beta(s,t) \sum_{i \in a_1} \sum_{j \in a_2} r_{i,j}(\underset{\sim}{x}) \\ &\geq st\, \beta(s,t)\, r_{\{1\},\{k\}}(\underset{\sim}{x}) \qquad \text{(Using Lemma 2.3.1)} \\ &\geq \beta(1,1) r_{\{1\},\{k\}}(\underset{\sim}{x}) \\ &= r_1(\underset{\sim}{x}, (\{\{1\}\}, \{\{k\}\})) \; \blacksquare \end{aligned}$$

Suppose $Y_i = X_{[i]}$, $\underset{\sim}{Y} = (Y_1, \ldots, Y_k) \in \mathscr{Y} = \{\underset{\sim}{y} : y_1 \leq y_2 \leq \ldots \leq y_k\}$. Then, as a consequence of Lemma 2.3.1 and Lemma 2.3.2, we have

Theorem 2.3.1 : Suppose assumptions of Lemma 2.3.2 are satisfied. Then $r(\tau, \delta)$ is minimized in $\mathscr{D}$ by δ_τ given by

$$\delta_\tau((\{g1\}, \{gk\}) | \underset{\sim}{y}) = \begin{cases} \dfrac{1}{|G_{\underset{\sim}{y}}|} & , \forall\, g \in G_{\underset{\sim}{y}} \, , \, \underset{\sim}{y} \in \mathscr{Y} \\ 0 & , \text{otherwise} \end{cases} \tag{2.3.3}$$

where $G_{\underset{\sim}{y}} = \{g \in G : g\underset{\sim}{y} = \underset{\sim}{y}\}$.

Proof : By Lemma 2.3.2, it follows that $(a_1, a_2) = (\{\{1\}\}, \{\{k\}\})$ minimizes $r_1(\underset{\sim}{x}, (a_1, a_2))$ over $(a_1, a_2) \in \mathscr{A}$. Hence, if $\underset{\sim}{X} = \underset{\sim}{y}$ is observed then an application of Lemma 2.2.1 proves the result. $\blacksquare$

Remark : For $\underset{\sim}{x} \in \mathscr{X}$, δ_τ is defined as :

$$\delta_\tau((a_1, a_2) | \underset{\sim}{x}) = \delta_\tau((ga_1, ga_2) | \underset{\sim}{y}) \text{ if and only if } g\underset{\sim}{x} = \underset{\sim}{y}.$$

Let

$$R(\underset{\sim}{\theta},\delta) = \int_{\mathcal{X}} \left\{ \sum_{(a_1,a_2)\in\mathcal{A}} \delta((a_1,a_2)|\underset{\sim}{x})\ L(\underset{\sim}{\theta},(a_1,a_2))\right\} \prod_{i=1}^{k} f(x_i;\theta_i)d\tau(\underset{\sim}{\theta})\ d\underset{\sim}{x}$$

denote the risk when decision rule $\delta \in \mathcal{D}$ is used and $\underset{\sim}{\theta} \in \Omega$ is the true value of the parameter. Also, let

$$\mathcal{D}_I = \{\delta \in \mathcal{D} : \delta((a_1,a_2)|\underset{\sim}{x}) = \delta((ga_1,ga_2)|g\underset{\sim}{x}),$$
$$\forall\ (a_1,a_2) \in \mathcal{A},\ \underset{\sim}{x} \in \mathcal{X} \text{ and } g \in G\}$$

denote the class of all permutation invariant decision rules.

<u>**Lemma 2.3.3**</u> **:** Suppose that the loss function L is permutation invariant. Then

$$R(g\underset{\sim}{\theta},\delta) = R(\underset{\sim}{\theta},\delta)$$

for every $\underset{\sim}{\theta} \in \Omega$, $g \in G$, and $\delta \in \mathcal{D}_I$.

<u>**Proof**</u> **:**

$$R(g\underset{\sim}{\theta},\delta) = \int_{\mathcal{X}} \left\{ \sum_{(a_1,a_2)\in\mathcal{A}} \delta((a_1,a_2)|\underset{\sim}{x})L(g\underset{\sim}{\theta},(a_1,a_2))\right\} \prod_{i=1}^{k} f(x_i;\theta_{g^{-1}i})\ d\underset{\sim}{x}$$

$$= \int_{\mathcal{X}} \left\{ \sum_{(a_1,a_2)\in\mathcal{A}} \delta((a_1,a_2)|\underset{\sim}{x})L(\underset{\sim}{\theta},(g^{-1}a_1,g^{-1}a_2))\right\} \prod_{i=1}^{k} f(x_i;\theta_{g^{-1}i})\ d\underset{\sim}{x}$$

(Since L is permutation invariant)

$$= \int_{\mathscr{X}} \left\{ \sum_{(a_1,a_2)\in\mathscr{A}} \delta((g^{-1}a_1, g^{-1}a_2) | g^{-1}\underset{\sim}{x}) L(\underset{\sim}{\theta}, (g^{-1}a_1, g^{-1}a_2)) \right\}$$

$$\prod_{i=1}^{k} f(x_i; \theta_{g^{-1}i})\, d\underset{\sim}{x}$$

(Since $\delta \in \mathscr{D}_I$)

$$= \int_{\mathscr{X}} \left\{ \sum_{(a_1,a_2)\in\mathscr{A}} \delta((a_1,a_2) | g^{-1}\underset{\sim}{x}) L(\underset{\sim}{\theta}, (a_1,a_2)) \right\}$$

$$\prod_{i=1}^{k} f(x_{gi}; \theta_i) d\underset{\sim}{x}$$

$$= \int_{\mathscr{X}} \left\{ \sum_{(a_1,a_2)\in\mathscr{A}} \delta((a_1,a_2) | \underset{\sim}{x}) L(\underset{\sim}{\theta}, (a_1,a_2)) \right\} \prod_{i=1}^{k} f(x_i; \theta_i) d\underset{\sim}{x}$$

$$= R(\underset{\sim}{\theta}, \delta). \quad \blacksquare$$

Theorem 2.3.2 : Suppose that the pdf's $f(.;\theta_i)$, $i = 1,\dots,k$ have MLR property and loss function is given by (2.3.2) with $\ell_{i,j}$'s satisfying properties (i) - (iv) of (2.3.1). Then for all $\underset{\sim}{\theta} \in \Omega$ and $\delta \in \mathscr{D}_I$

$$R(\underset{\sim}{\theta}, \delta_\tau) \le R(\underset{\sim}{\theta}, \delta)$$

where δ_τ is defined by (2.3.3).

Proof : Clearly $\delta_\tau \in \mathscr{D}_I$. Fix $\underset{\sim}{\theta} \in \Omega$ and let τ be the probability distribution on Ω which puts mass $\frac{1}{\lfloor k}$ on each permutation of fixed $\underset{\sim}{\theta} \in \Omega$. Then clearly τ is symmetric and hence from Theorem 2.3.1, we have

$$r(\tau,\delta_\tau) \le r(\tau,\delta), \qquad \forall\ \delta \in \mathcal{D}$$

However, if $\delta \in \mathcal{D}_I$, then

$$r(\tau,\delta) = \frac{1}{\lfloor k} \sum_{g\in G} R(g\underset{\sim}{\theta},\delta)$$

$$= R(\underset{\sim}{\theta},\delta) \qquad \text{(using Lemma 2.3.3).}$$

Hence,

$$R(\underset{\sim}{\theta},\delta_\tau) = r(\tau,\delta_\tau) \le r(\tau,\delta) = R(\underset{\sim}{\theta},\delta), \quad \forall\ \delta \in \mathcal{D}_I \ \blacksquare$$

From the above theorem we conclude that δ_τ is both minimax and admissible with in the class $\mathcal{D}_I$ of permutation invariant decision rules. Since the group of permutations G is finite, it follows that δ_τ is both minimax and admissible in $\mathcal{D}$.

2.4 Bayes Rules When the Selected Subsets are Required to be Disjoint: Specific Loss Function

In case there seems to be strong evidence that the LEP and the UEP are 'distinctly' o
one may like at the selected subsets should be disjoint. Thus, we now

~~Consider~~ the goal of simultaneously selecting two non-empty and disjoint subsets S_L and S_U containing LEP and UEP respectively, using Bayesian approach. For the above goal our action space $\mathcal{A}_d$ is given by

$$\mathcal{A}_d = \{(a_1,a_2) : a_1 \subset \{1,\ldots,k\},\ a_2 \subset \{1,\ldots,k\},$$
$$a_1 \ne \phi,\ a_2 \ne \phi,\ \ a_1 \cap a_2 = \phi\},$$

which contains $3^k - 2^{k+1}+1$ elements. Suppose that the loss function is given by

$$L(\underset{\sim}{\theta},(a_1,a_2)) = c_1|a_1|+c_2|a_2|+c_3(\min_{i\in a_1} \theta_i - \theta_{[1]})$$
$$+c_4(\theta_{[k]}-\max_{i\in a_2} \theta_i) \qquad (2.4.1)$$

where $c_i \ge 0$, $i = 1,2,3,4$, are some constants whose relative

values indicate the relative weights given to the four components of the loss function.

Since we are interested in finding Bayes rule, it is sufficient to consider only the non-randomized decision rules. The following lemma reduces the number of decision rules to be compared from $3^k - 2^{k+1} + 1$ to $\frac{k(k-1)}{2}$.

Lemma 2.4.1 : Suppose that the pdf's $f(.;\theta_i)$, $i = 1,\dots,k$ have the MLR property and the prior distribution τ is symmetric on Ω. Then

$$\min_{\substack{1\le s,t\le k-1 \\ 2\le s+t\le k}} r(\tau,\delta^{s,t}) = \min_{\delta\in D} r(\tau,\delta)$$

where $\delta^{s,t}$ $(1 \le s,t \le k-1,\ 2 \le s+t \le k)$ is given by

$$\delta^{s,t}((a_1,a_2)|\underset{\sim}{x}) = \begin{cases} 1 & , \text{ if } a_1 = a_{\{s\}} \text{ and } a_2 = a^{\{t\}} \\ & \\ 0 & , \quad \text{otherwise} \end{cases}$$

Proof : Consider (s,t)-th decision problem where action space is

$$\mathcal{A}_d(s,t) = \{(a_1,a_2) : (a_1,a_2) \in \mathcal{A}_d, |a_1| = s, |a_2| = t\},$$
$$1 \le s,t \le k-1,\ 2 \le s+t \le k.$$

For $s+t < k$ with the observation vector $\underset{\sim}{x}$, the action space $\mathcal{A}_d(s,t)$ and the loss function given by (2.4.1), (s,t)-th decision problem is equivalent to partitioning the set $\{1,\dots,k\}$ into three disjoint subsets γ_1, γ_2, and γ_3, where γ_1 is of size t, γ_3 is of size s and γ_2 is of size k-s-t. Conditions of Eaton's (1967) paper are satisfied and thus the rule which assigns $\{\{1\},\dots,\{s\}\}$ to γ_3, $\{\{k-t+1\},\dots,\{k\}\}$ to γ_1 and $\{\{s+1\},\dots,\{k-t\}\}$ to γ_2 is Bayes.

For $s+t = k$, (s,t)-th decision problem with the observation vector $\underset{\sim}{x}$, the action space $\mathcal{A}_d(s,t)$ and the loss function given by (2.4.1) is equivalent to partitioning the set $\{1,\ldots,k\}$ into two disjoint subsets γ_1 and γ_3, where γ_1 is of size t and γ_3 is of size $s = k-t$. Again the results of Eaton's (1967) paper are applicable. Hence the rule which assigns $\{\{1\},\ldots,\{s\}\}$ to γ_3 and $\{\{s+1\},\ldots,\{k\}\}$ to γ_1 is Bayes. Thus

$$r(\tau,\delta^{s,t}) \le r(\tau,\delta), \qquad \text{for all } \delta \in D(s,t)$$

where $D(s,t)$ is the set of all non-randomized decision rules corresponding to the action space $\mathcal{A}_d(s,t)$. Hence Bayes rule is one of the decision rules $\delta^{s,t}$, $1 \le s,t \le k-1$, $2 \le s+t \le k$. ■

Remark : Lemma 2.4.1 holds for any loss function which satisfies the conditions of Eaton's (1967) paper.

Define

$$\Delta_s(t) = r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{t+1\}})) - r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{t\}})),$$

$$t = 1,\ldots,k-2,\ 2 \le s+t \le k-1$$

and

$$\Delta_t^*(s) = r_1(\underset{\sim}{x},(a_{\{s+1\}},a^{\{t\}})) - r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{t\}})),$$

$$s = 1,\ldots,k-2,\ 2 \le s+t \le k.$$

The following Lemmas 2.4.2 and 2.4.3 further reduce the number of comparisions to be made for finding Bayes rule.

Lemma 2.4.2 : Under the assumptions of Lemma 2.4.1,

(i) $\Delta_s(t) \ge \Delta_s(t-1)$, $\quad t = 2,\ldots,k-2,\ 2 \le s+t \le k$

(ii) $\Delta_t^*(s) \ge \Delta_t^*(s-1)$, $\quad s = 2,\ldots,k-2,\ 2 \le s+t \le k.$

<u>**Proof**</u> : (i) $\Delta_s(t) = r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{t+1\}})) - r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{t\}}))$.

Let,

$$b_s = \min\{\theta_{\{1\}},\dots,\theta_{\{s\}}\},$$
$$b^t = \max\{\theta_{\{k-t+1\}},\dots,\theta_{\{k\}}\},$$
$$A_1 = \{\underset{\sim}{\theta} : \theta_{\{k-t\}} > b^t\}$$
$$A_2 = \{\underset{\sim}{\theta} : \theta_{\{k-t+1\}} > b^{t-1}\}$$
$$A_3 = \{\underset{\sim}{\theta} : \theta_{\{k-t\}} > b^{t-1}\}$$

and

$$B = \{\underset{\sim}{\theta} : \theta_{\{k-t+1\}} > \theta_{\{k-t\}}\}.$$

Then

$$\Delta_s(t) = \int_\Omega (c_1 s + c_2(t+1) + c_3(b_s - \theta_{[1]}) + c_4(\theta_{[k]} - b^{t+1})) \prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta})$$
$$-\int_\Omega (c_1 s + c_2 t + c_3(b_s - \theta_{[1]}) + c_4(\theta_{[k]} - b^{t})) \prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta})$$
$$= c_2 \int_\Omega \prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta}) - c_4 \int_\Omega (b^{t+1}-b^t)\prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta})$$
$$= c_2 \int_\Omega \prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta}) - c_4 \int_{A_1} (\theta_{\{k-t\}}-b^t)\prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta})$$

(Since outside A_1, $b^{t+1} = b^t$)

$$\Delta_s(t) - \Delta_s(t-1) = c_4 \int_{A_2} (\theta_{\{k-t+1\}} - b^{t-1}) \prod_{l=1}^{k} f(x_l;\theta_l)\, d\tau(\underset{\sim}{\theta})$$
$$- c_4 \int_{A_1} (\theta_{\{k-t\}} - b^{t}) \prod_{l=1}^{k} f(x_l;\theta_l)\, d\tau(\underset{\sim}{\theta})$$

$$\geq c_4 \int_{A_2} (\theta_{\{k-t+1\}} - b^{t-1}) \prod_{l=1}^{k} f(x_l;\theta_l)\, d\tau(\underset{\sim}{\theta})$$

$$- c_4 \int_{A_1} (\theta_{\{k-t\}} - b^{t-1}) \prod_{l=1}^{k} f(x_l;\theta_l)\, d\tau(\underset{\sim}{\theta})$$

(Since $b^{t-1} \leq b^{t}$)

$$\geq c_4 \int_{A_2} (\theta_{\{k-t+1\}} - b^{t-1}) \prod_{l=1}^{k} f(x_l;\theta_l)\, d\tau(\underset{\sim}{\theta})$$

$$- c_4 \int_{A_3} (\theta_{\{k-t\}} - b^{t-1}) \prod_{l=1}^{k} f(x_l;\theta_l)\, d\tau(\underset{\sim}{\theta}).$$

(Since $A_1 \subset A_3$)

$$= c_4 \Bigg[\int_{A_2 \cap B} (\theta_{\{k-t+1\}} - b^{t-1}) \prod_{l=1}^{k} f(x_l;\theta_l)\, d\tau(\underset{\sim}{\theta})$$

$$+ \int_{A_2 \cap B^c} (\theta_{\{k-t+1\}} - b^{t-1}) \prod_{l=1}^{k} f(x_l;\theta_l)\, d\tau(\underset{\sim}{\theta})$$

$$- \int_{A_3 \cap B} (\theta_{\{k-t\}} - b^{t-1}) \prod_{l=1}^{k} f(x_l;\theta_l)\, d\tau(\underset{\sim}{\theta})$$

$$- \int_{A_3 \cap B^c} (\theta_{\{k-t\}} - b^{t-1}) \prod_{l=1}^{k} f(x_l;\theta_l)\, d\tau(\underset{\sim}{\theta}) \Bigg]$$

On interchanging the roles of $\theta_{\{k-t\}}$ and $\theta_{\{k-t+1\}}$ in 2nd and 4th integral, we get

$$\Delta_s(t) - \Delta_s(t-1)$$

$$\geq c_4 \int_{A_2 \cap B} (\theta_{\{k-t+1\}} - b^{t-1}) \prod_{\substack{l=1 \\ l \neq k-t, k-t+1}}^{k} f(x_{[l]};\theta_{\{l\}})$$

$$[f(x_{[k-t]};\theta_{\{k-t\}}) f(x_{[k-t+1]};\theta_{\{k-t+1\}})$$

$$-f(x_{[k-t]};\theta_{\{k-t+1\}}) f(x_{[k-t+1]};\theta_{\{k-t\}})] d\tau(\underset{\sim}{\theta})$$

$$- c_4 \int_{A_3 \cap B} (\theta_{\{k-t\}} - b^{t-1}) \prod_{\substack{l=1 \\ l \neq k-t, k-t+1}}^{k} f(x_{[l]};\theta_{\{l\}})$$

$$[f(x_{[k-t]};\theta_{\{k-t\}}) f(x_{[k-t+1]};\theta_{\{k-t+1\}})$$

$$-f(x_{[k-t]};\theta_{\{k-t+1\}}) f(x_{[k-t+1]};\theta_{\{k-t\}})] d\tau(\underset{\sim}{\theta})$$

$$\geq c_4 \int_{A_2 \cap B} (\theta_{\{k-t+1\}} - b^{t-1}) \prod_{\substack{l=1 \\ l \neq k-t, k-t+1}}^{k} f(x_{[l]};\theta_{\{l\}})$$

$$[f(x_{[k-t]};\theta_{\{k-t\}}) f(x_{[k-t+1]};\theta_{\{k-t+1\}})$$

$$-f(x_{[k-t]};\theta_{\{k-t+1\}}) f(x_{[k-t+1]};\theta_{\{k-t\}})] d\tau(\underset{\sim}{\theta})$$

$$- c_4 \int_{A_2 \cap B} (\theta_{\{k-t\}} - b^{t-1}) \prod_{\substack{l=1 \\ l \neq k-t, k-t+1}}^{k} f(x_{[l]};\theta_{\{l\}})$$

$$[f(x_{[k-t]};\theta_{\{k-t\}}) f(x_{[k-t+1]};\theta_{\{k-t+1\}})$$

$$-f(x_{[k-t]};\theta_{\{k-t+1\}}) f(x_{[k-t+1]};\theta_{\{k-t\}})] d\tau(\underset{\sim}{\theta})$$

$$(\text{Since } A_3 \cap B \subset A_2 \cap B)$$

$$= c_4 \int_{A_2 \cap B} (\theta_{\{k-t+1\}} - \theta_{\{k-t\}}) \prod_{\substack{l=1 \\ l \neq k-t, k-t+1}}^{k} f(x_{[l]};\theta_{\{l\}})$$

$$[f(x_{[k-t]};\theta_{\{k-t\}}) f(x_{[k-t+1]};\theta_{\{k-t+1\}})$$

$$-f(x_{[k-t]};\theta_{\{k-t+1\}}) f(x_{[k-t+1]};\theta_{\{k-t\}})] d\tau(\underset{\sim}{\theta})$$

Now using MLR property and the fact that $\theta_{\{k-t+1\}} > \theta_{\{k-t\}}$ for $\underset{\sim}{\theta} \in B$, we get

$$\Delta_s(t) - \Delta_s(t-1) \geq 0$$

(ii) Let

$$B_1 = \{\underset{\sim}{\theta} : \theta_{\{s+1\}} < b_s\}$$

$$B_2 = \{\underset{\sim}{\theta} : \theta_{\{s\}} < b_{s-1}\}$$

$$B_3 = \{\underset{\sim}{\theta} : \theta_{\{s+1\}} < b_{s-1}\}$$

and

$$D = \{\underset{\sim}{\theta} : \theta_{\{s+1\}} > \theta_{\{s\}}\}.$$

Then,

$$\begin{aligned}
\Delta_t^*(s) &= r_1(\underset{\sim}{x},(a_{\{s+1\}},a^{\{t\}})) - r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{t\}})) \\
&= c_1 \int_\Omega \prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta}) + c_3 \int_\Omega (b_{s+1}-b_s) \prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta}) \\
&= c_1 \int_\Omega \prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta}) + c_3 \int_B (\theta_{\{s+1\}}-b_s) \prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta})
\end{aligned}$$

(Since outside $B_1, b_{s+1} = b_s$)

Therefore,

$$\begin{aligned}
\Delta_t^*(s)-\Delta_t^*(s-1) &= c_3 \int_{B_1} (\theta_{\{s+1\}}-b_s) \prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta}) \\
&\quad - c_3 \int_{B_2} (\theta_{\{s\}}-b_{s-1}) \prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta}) \\
&\geq c_3 \int_{B_1} (\theta_{\{s+1\}}-b_{s-1}) \prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta}) \\
&\quad - c_3 \int_{B_2} (\theta_{\{s\}}-b_{s-1}) \prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta})
\end{aligned}$$

(Since $b_{s-1} \geq b_s$)

$$\geq c_3 \int_{B_3} (\theta_{\{s+1\}}-b_{s-1}) \prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta})$$

$$- c_3 \int_{B_2} (\theta_{\{s\}}-b_{s-1}) \prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta})$$

(Since $B_1 \subset B_3$ and $\theta_{\{s+1\}} < b_{s-1}$ in B_1 and B_3)

$$= c_3 \Big[\int_{B_3\cap D} (\theta_{\{s+1\}}-b_{s-1}) \prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta})$$

$$+ \int_{B_3\cap D^c} (\theta_{\{s+1\}}-b_{s-1}) \prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta})$$

$$- \int_{B_2\cap D} (\theta_{\{s\}}-b_{s-1}) \prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta})$$

$$- \int_{B_2\cap D^c} (\theta_{\{s\}}-b_{s-1}) \prod_{l=1}^{k} f(x_l;\theta_l)d\tau(\underset{\sim}{\theta}) \Big]$$

On interchanging the roles of $\theta_{\{s\}}$ and $\theta_{\{s+1\}}$ in 2nd and 4th integral, we get

$$\Delta_t^*(s) - \Delta_t^*(s-1)$$

$$\geq c_3 \int_{B_3\cap D} (\theta_{\{s+1\}}-b_{s-1}) \prod_{\substack{l=1 \\ l\neq s,s+1}}^{k} f(x_{[l]};\theta_{\{l\}})$$

$$[f(x_{[s]};\theta_{\{s\}})f(x_{[s+1]};\theta_{\{s+1\}})$$

$$-f(x_{[s]};\theta_{\{s+1\}})f(x_{[s+1]};\theta_{\{s\}})]d\tau(\underset{\sim}{\theta})$$

$$- c_3 \int_{B_2\cap D} (\theta_{\{s\}}-b_{s-1}) \prod_{\substack{l=1 \\ l\neq s,s+1}}^{k} f(x_{[l]};\theta_{\{l\}})$$

$$[f(x_{[s]};\theta_{\{s\}})f(x_{[s+1]};\theta_{\{s+1\}})$$

$$-f(x_{[s]};\theta_{\{s+1\}})f(x_{[s+1]};\theta_{\{s\}})]d\tau(\underset{\sim}{\theta})$$

$$\geq c_3 \int_{B_2 \cap D} (\theta_{\{s+1\}} - b_{s-1}) \prod_{\substack{l=1 \\ l \neq s, s+1}}^{k} f(x_{[l]};\theta_{\{l\}})$$

$$[f(x_{[s]};\theta_{\{s\}}) f(x_{[s+1]};\theta_{\{s+1\}})$$

$$-f(x_{[s]};\theta_{\{s+1\}}) f(x_{[s+1]};\theta_{\{s\}})] d\tau(\underset{\sim}{\theta})$$

$$- c_3 \int_{B_2 \cap D} (\theta_{\{s\}} - b_{s-1}) \prod_{\substack{l=1 \\ l \neq s, s+1}}^{k} f(x_{[l]};\theta_{\{l\}})$$

$$[f(x_{[s]};\theta_{\{s\}}) f(x_{[s+1]};\theta_{\{s+1\}})$$

$$-f(x_{[s]};\theta_{\{s+1\}}) f(x_{[s+1]};\theta_{\{s\}})] d\tau(\underset{\sim}{\theta})$$

(Since $B_2 \cap D \subset B_3 \cap D$ and $b_{s-1} > \theta_{\{s\}}$ in $B_2 \cap$ D and $B_3 \cap D$). Now using MLR property and the fact that $\theta_{\{s+1\}} < \theta_{\{s\}}$ in $B_2 \cap D$, we get

$$\Delta_t^*(s) - \Delta_t^*(s-1) \geq 0 \quad ■$$

<u>**Note**</u> **:** $\Delta_s(t)$ is independent of s and $\Delta_t^*(s)$ is independent of t. In particular,

$$\Delta_s(t) = \Delta_{s+1}(t) \quad , \quad 1 \leq s \leq k-2,\ 2 \leq s+t \leq k-1$$

and

$$\Delta_t^*(s) = \Delta_{t+1}^*(s) \quad , \quad 1 \leq t \leq k-2,\ 2 \leq s+t \leq k-1.$$

Let

$$F_1 = \{t : \Delta_1(t) \geq 0\}$$

and

$$F_2 = \{s : \Delta_1^*(s) \geq 0\}.$$

For $F_1 \neq \phi$, and $F_2 \neq \phi$, let m and p denote the smallest integers in F_1 and F_2 respectively. Following theorem further reduces the number of integrals to be evaluated for obtaining Bayes rules.

<u>**Theorem 2.4.1**</u> **:** Under the assumptions of Lemma 2.4.1, the Bayes rule selects subset pair (a_1^*, a_2^*) with probability one, where

$$
\begin{aligned}
r_1(\underset{\sim}{x},(a_1^*,a_2^*)) &= \min_{1\le i\le k-1} r_1(\underset{\sim}{x},(a_{\{i\}},a^{\{k-i\}})), \text{ if } F_1 = \phi,\ F_2 = \phi \\
&= \min_{k-m\le i\le k-1} r_1(\underset{\sim}{x},(a_{\{i\}},a^{\{k-i\}})), \text{ if } F_1 \ne \phi,\ F_2 = \phi \\
&= \min_{1\le i\le p} r_1(\underset{\sim}{x},(a_{\{i\}},a^{\{k-i\}})), \text{ if } F_1 = \phi,\ F_2 \ne \phi \\
&= r_1(\underset{\sim}{x},(a_{\{p\}},a^{\{m\}})), \text{ if } F_1 \ne \phi,\ F_2 \ne \phi, \text{ and } m+p \le k \\
&= \min_{k-m\le i\le p} r_1(\underset{\sim}{x},(a_{\{i\}},a^{\{k-i\}})), \text{ if } F_1 \ne \phi,\ F_2 \ne \phi \\
&\qquad\qquad \text{and } m+p \ge k+1
\end{aligned}
\tag{2.4.2}
$$

<u>**Proof**</u> **:** In view of essential completeness of non-randomized decision rules, it is enough to show that

$$r_1(\underset{\sim}{x},(a_1^*,a_2^*)) = \min_{(a_1,a_2)\in\mathscr{A}} r_1(\underset{\sim}{x},(a_1,a_2))$$

<u>**Case I**</u> **:** $F_1 = \phi,\ F_2 = \phi$

$$
\begin{aligned}
F_1 = \phi &\implies \Delta_1(t) < 0,\ t = 1,\dots,k-1 \\
&\implies \Delta_s(t) < 0,\ t = 1,\dots,k-s-1,\ s = 1,\dots,k-1 \\
&\qquad\qquad (\text{Since } \Delta_s(t) = \Delta_{s+1}(t)) \\
&\implies r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{1\}})) > r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{2\}})) > \dots \\
&\qquad > r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{k-s\}})), \qquad s = 1,\dots,k-1
\end{aligned}
$$

Similarly $F_2 = \phi$, implies that

$$
\begin{aligned}
&r_1(\underset{\sim}{x},(a_{\{1\}},a^{\{t\}})) > r_1(\underset{\sim}{x},(a_{\{2\}},a^{\{t\}})) > \dots \\
&\qquad > r_1(\underset{\sim}{x},(a_{\{k-t\}},a^{\{t\}})), \qquad t = 1,\dots,k-1
\end{aligned}
$$

Therefore $F_1 = \phi$, $F_2 = \phi$ implies

$$r_1(\underset{\sim}{x},(a_1^*,a_2^*)) = \min_{(a_1,a_2)\in\mathscr{A}} r_1(\underset{\sim}{x},(a_1,a_2))$$

$$= \min_{1\le i\le k-1} r_1(\underset{\sim}{x},(a_{\{i\}},a^{\{k-i\}}))$$

Case II : $F_1 \neq \phi$, $F_2 = \phi$, m being the smallest integer in F_1

$F_1 \neq \phi$ implies that

$\Delta_1(t) < 0$, $t = 1,\ldots,m-1$, and $\Delta_1(t) \ge 0$, $t = m,\ldots,k-2$

which in turn implies that for $s = 1,\ldots,k-m-1$

$\Delta_s(t) < 0$, $t = 1,\ldots,m-1$, and $\Delta_s(t) \ge 0$, $t = m,\ldots,k-s-1$

Therefore,

$$r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{1\}})) > r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{2\}})) >\ldots> r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{m\}}))$$

and

$$r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{m\}})) > r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{m+1\}})) > \ldots$$

$$> r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{k-s\}})), \qquad \text{for } s = 1,\ldots,k-m-1$$

But, since $F_2 = \phi$, we have

$$r_1(\underset{\sim}{x},(a_{\{1\}},a^{\{t\}})) > \ldots > r_1(\underset{\sim}{x},(a_{\{k-t\}},a^{\{t\}})), \quad t = 1,\ldots,k-1.$$

Hence,

$$r_1(\underset{\sim}{x},(a_1^*,a_2^*)) = \min_{(a_1,a_2)\in\mathscr{a}} r_1(\underset{\sim}{x},(a_1,a_2))$$

$$= \min_{k-m\le i\le k-1} r_1(\underset{\sim}{x},(a_{\{i\}},a^{\{k-i\}}))$$

Case III : $F_1 = \phi$, $F_2 \neq \phi$, p being the smallest integer in F_2.

Now $F_2 \neq \phi$, implies that

$$r_1(\underset{\sim}{x},(a_{\{1\}},a^{\{t\}})) > r_1(\underset{\sim}{x},(a_{\{2\}},a^{\{t\}})) >\ldots> r_1(\underset{\sim}{x},(a_{\{p\}},a^{\{t\}}))$$

and

$$r_1(\underset{\sim}{x},(a_{\{p\}},a^{\{t\}})) < r_1(\underset{\sim}{x},(a_{\{p+1\}},a^{\{t\}})) < \ldots$$

$$< r_1(\underset{\sim}{x},(a_{\{k-t\}},a^{\{t\}})), \qquad \text{for } t = 1,\ldots,k-p-1.$$

But since $F_1 = \phi$, we have

$$r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{1\}})) > \dots > r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{k-s\}})),\ s = 1,\dots,k-1.$$

Hence,

$$r_1(\underset{\sim}{x},(a_1^*,a_2^*)) = \min_{(a_2,a_2)\in \mathcal{a}} r_1(\underset{\sim}{x},(a_1,a_2))$$

$$= \min_{1\le i\le p} r_1(\underset{\sim}{x},(a_{\{i\}},a^{\{k-1\}}))$$

Case IV : $F_1 \ne \phi$, $f_2 \ne \phi$, m and p being smallest integers in F_1 and F_2, respectively

$$F_1 \ne \phi \Longrightarrow \begin{cases} r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{1\}})) > \dots > r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{m\}})) \\ \text{and} \\ r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{m\}})) < \dots < r_1(\underset{\sim}{x},(a_{\{s\}},a^{\{k-s\}})) \end{cases}$$

for $s = 1,\dots,k-m-1$

$$F_2 \ne \phi \Longrightarrow \begin{cases} r_1(\underset{\sim}{x},(a_{\{1\}},a^{\{t\}})) > \dots > r_1(\underset{\sim}{x},(a_{\{p\}},a^{\{t\}})) \\ \text{and} \\ r_1(\underset{\sim}{x},(a_{\{p\}},a^{\{t\}})) < \dots < r_1(\underset{\sim}{x},(a_{\{k-t\}},a^{\{t\}})) \end{cases}$$

for $t = 1,\dots,k-p-1$

Therefore,

$$r_1(\underset{\sim}{x},(a_1^*,a_2^*)) = \min_{(a_1,a_2)\in \mathcal{a}} r_1(\underset{\sim}{x},(a_1,a_2))$$

$$= \begin{cases} r_1(\underset{\sim}{x},(a_{\{p\}},a^{\{m\}})), & \text{if } m+p \le k \\ \\ \displaystyle\min_{k-m\le i\le p} r_1(\underset{\sim}{x},(a_{\{i\}},a^{\{k-1\}})), & \text{if } m+p \ge k+1 \end{cases}$$

■

Note : Goel and Rubin (1977) dealt with the problem of selecting a non-empty subset containing UEP, using Bayesian approach. For the loss function given by

$$L_1(\underset{\sim}{\theta},a) = c_2|a| + c_4(\theta_{[k]} - \max_{i\in a} \theta_i)$$

(where $c_2 \geq 0$, $c_4 \geq 0$, and $a \subset \{1,\ldots,k\}$, $a \neq \phi$) they derived non-randomized Bayes rules. Similarly one can obtain Bayes rules for selecting a non-empty subset containing LEP when the loss function is given by

$$L_2(\underset{\sim}{\theta},a) = c_1\ |a| + c_3\ (\min_{i\in a} \theta_i - \theta_{[1]})$$

where $c_1 \geq 0$, $c_2 \geq 0$, $a \subset \{1,\ldots,k\}$, $a \neq \phi$. Note that for the loss function $L(\underset{\sim}{\theta},(a_1,a_2)) = L_1(\underset{\sim}{\theta},a_1) + L_2(\underset{\sim}{\theta},a_2)$, the separate solutions of above two problems give non-randomized Bayes rules for the goal of simultaneously selecting two non-empty subsets (not necessarily disjoint) containing LEP and UEP. Therefore, with the specific loss function considered in this section, if the subsets are not required to be disjoint, the trivial solution is to use Goel and Rubin (1977) procedure separately for selecting LEP and UEP.

2.5 Application to Normal Distribution

Suppose that, given $\theta_1,\ldots,\theta_k$, the random variables $X_1,\ldots,X_k$ are independently normally distributed with means $\theta_1,\ldots,\theta_k$, respectively, and a common known variance σ^2. The unknown and/or unequal variances case is a topic of future research. Further assume that $\theta_1,\ldots,\theta_k$ are independently and identically distributed normal random variables with mean μ and variance ξ^2. Thus, here we have

$$f(x_i,\theta_i) = \frac{1}{\sigma}\,\phi\left(\frac{x_i-\theta_i}{\sigma}\right),\quad i = 1,\ldots,k$$

$$d\tau(\underset{\sim}{\theta}) = \frac{1}{\xi^n}\prod_{i=1}^{k}\phi\left(\frac{\theta_i-\mu}{\xi}\right)\,d\underset{\sim}{\theta}$$

and $\Omega = \mathcal{X} = \mathbb{R}^k$, the k-dimensional Euclidean space. Let $m(\underset{\sim}{x})$ denote the marginal pdf of $\underset{\sim}{X} = (X_1,\ldots,X_k)$, that is

$$m(\underset{\sim}{x}) = \int_{\Omega} \prod_{i=1}^{k} f(x_i;\theta_i)\, d\tau(\underset{\sim}{\theta})$$

so that,

$$\frac{\Delta_s(t)}{m(\underset{\sim}{x})} = c_2 - c_4\, E((\theta_{\{k-t\}} - b^t)^+ | \underset{\sim}{X} = \underset{\sim}{x}), \qquad t = 1,\ldots,k-2$$

$$2 \le s+t \le k$$

and

$$\frac{\Delta_t^*(s)}{m(\underset{\sim}{x})} = c_1 - c_3\, E((b_s - \theta_{\{s+1\}})^+ | \underset{\sim}{X} = \underset{\sim}{x}), \qquad s = 1,\ldots,k-2$$

$$2 \le s+t \le k$$

where,

$$z^+ = \begin{cases} z, & \text{if } z \ge 0 \\ 0, & \text{otherwise.} \end{cases}$$

Now, given $\underset{\sim}{X} = \underset{\sim}{x}$, the random variables $\theta_{\{1\}},\ldots,\theta_{\{k\}}$ are independently distributed with $\theta_{\{i\}}$ $(1 \le i \le k)$ having a normal distribution with mean $\alpha_i = \left[\frac{x_{[i]}}{\sigma^2} + \frac{\mu}{\xi^2}\right]\eta^2$ and the variance $\eta^2 = \left[\frac{1}{\sigma^2} + \frac{1}{\xi^2}\right]^{-1}$.

<u>**Lemma 2.5.1**</u> **:** Under the above set-up

(i) $$\frac{\Delta_s(t)}{m(\underset{\sim}{x})} = c_2 - c_4\eta \int_{-\infty}^{\infty} \prod_{i=k-t+1}^{k} \Phi\left(z + \frac{\alpha_{k-t} - \alpha_i}{\eta}\right) \Phi(-z)\,dz$$

$$t = 1,\ldots,k-2,\ 2 \le s+t \le k$$

(ii) $$\frac{\Delta_s^*(t)}{m(\underset{\sim}{x})} = c_1 - c_3\eta \int_{-\infty}^{\infty} \prod_{i=1}^{s} \Phi\left(z + \frac{\alpha_i - \alpha_{s+1}}{\eta}\right) \Phi(-z)\,dz$$

$$s = 1,\ldots,k-2,\ 2 \le s+t \le k$$

Proof : (i) Let $Z_1, \ldots, Z_k$ denote k independent standard normal variates. Then for $\lambda > 0$, we have

$$H(\lambda|\underset{\sim}{x}) = P((\theta_{\{k-t\}} - b^t)^+ > \lambda | \underset{\sim}{X} = \underset{\sim}{x})$$

$$= P(\theta_{\{k-t\}} > \lambda + \theta_{\{i\}},\ i = k-t+1, \ldots, k | \underset{\sim}{X} = \underset{\sim}{x})$$

$$= P(Z_{k-t}\,\eta + \alpha_{k-t} > \lambda + Z_i\,\eta + \alpha_i,\ i = k-t+1, \ldots, k)$$

$$= \int_{-\infty}^{\infty} \prod_{i=k-t+1}^{k} \Phi(z + \frac{\alpha_{k-t} - \alpha_i}{\eta})\ \phi(z + \frac{\lambda}{\eta})\ dz$$

We know that for a non-negative continuous random variable X with distribution function F_X

$$E(X) = \int_{-\infty}^{\infty} (1 - F_X(x))dx$$

Hence,

$$E((\theta_{\{k-t\}} - b^t)^+ | \underset{\sim}{X} = \underset{\sim}{x}) = \int_0^{\infty} H(\lambda|\underset{\sim}{x})\ d\lambda$$

$$= \eta \int_{-\infty}^{\infty} \prod_{i=k-t+1}^{k} \Phi(z + \frac{\alpha_{k-t} - \alpha_i}{\eta})\ \Phi(-z)\ dz$$

and

$$\frac{\Delta_s(t)}{m(\underset{\sim}{x})} = c_2 - c_4\eta \int_{-\infty}^{\infty} \prod_{i=k-t+1}^{k} \Phi(z + \frac{\alpha_{k-t} - \alpha_i}{\eta})\ \Phi(-z)dz$$

(ii) Application of (i) on $W_i = -X_i$ yields the required result. ■

Define,

$$J_m(z) = \int_{-\infty}^{\infty} \Phi^m(z+u)\ \Phi(-u)du$$

Note that $J_m(z)$ is an increasing function of z and a decreasing function of m. Also

$$J_1(0) = \int_{-\infty}^{\infty} \Phi(u)\, \Phi(-u)du$$

Integration by parts gives

$$J_1(0) = \frac{1}{\sqrt{\pi}}$$

Lemma 2.5.2 : For $t = 1,\ldots,k,\ 2 \le s+t \le k,$

$$c_2 - c_4\, \eta \min\left\{J_1\left[\frac{\alpha_{k-t}-\alpha_k}{\eta}\right],\ J_t\left[\frac{\alpha_{k-t}-\alpha_{k-t+1}}{\eta}\right]\right\}$$

$$\le \frac{\Delta_s(t)}{m(\underset{\sim}{x})} \le c_2 - c_4\, \eta\, J_t\left[\frac{\alpha_{k-t}-\alpha_k}{\eta}\right]$$

and

$$c_1 - c_3\, \eta \min\left\{J_1\left[\frac{\alpha_1-\alpha_{t+1}}{\eta}\right],\ J_t\left[\frac{\alpha_t-\alpha_{t+1}}{\eta}\right]\right\}$$

$$\le \frac{\Delta_s^*(t)}{m(\underset{\sim}{x})} \le c_1 - c_3\, \eta\, J_t\left[\frac{\alpha_1-\alpha_{t+1}}{\eta}\right].$$

Proof : $$\frac{\Delta_s(t)}{m(\underset{\sim}{x})} = c_2 - c_4\, E((\theta_{\{p-t\}}-b^t)^+ | \underset{\sim}{X} = \underset{\sim}{x})$$

$$\ge c_2 - c_4\, E((\theta_{\{k-t\}}-\theta_{\{k\}})^+ | \underset{\sim}{X} = \underset{\sim}{x})$$

$$= c_2 - c_4 \int_0^{\infty} P(\theta_{\{k-t\}} > \lambda + \theta_{\{k\}} | \underset{\sim}{X} = \underset{\sim}{x})\, d\lambda$$

$$= c_2 - c_4 \int_0^{\infty} \left\{\int_{-\infty}^{\infty} \Phi\left[z + \frac{\alpha_{k-t}-\alpha_k}{\eta}\right] \phi(z + \frac{\lambda}{\eta})dz\right\} d\lambda$$

$$= c_2 - c_4\, \eta \int_{-\infty}^{\infty} \Phi\left[z + \frac{\alpha_{k-t}-\alpha_k}{\eta}\right] \Phi(-z)dz$$

$$= c_2 - c_4\, \eta\, J_1\left[\frac{\alpha_{k-t}-\alpha_k}{\eta}\right] \tag{2.5.1}$$

Also from (i) of Lemma 2.5.1, we have

$$\frac{\Delta_s(t)}{m(\underset{\sim}{x})} = c_2 - c_4 \eta \int_{-\infty}^{\infty} \prod_{i=k-t+1}^{k} \Phi\left(z + \frac{\alpha_{k-t}-\alpha_i}{\eta}\right) \Phi(-z)dz$$

Therefore,

$$c_2 - c_4 \eta J_t\left[\frac{\alpha_{k-t}-\alpha_{k-t+1}}{\eta}\right] \le \frac{\Delta_s(t)}{m(\underset{\sim}{x})}$$

$$\le c_2 - c_4 \eta J_t \left[\frac{\alpha_{k-t}-\alpha_k}{\eta}\right] \qquad (2.5.2)$$

From (2.5.1), and (2.5.2), we obtain

$$c_2 - c_4 \eta \min \left[J_1 \left[\frac{\alpha_{k-t}-\alpha_k}{\eta}\right], J_t \left[\frac{\alpha_{k-t}-\alpha_{k-t+1}}{\eta}\right]\right]$$

$$\le \frac{\Delta_s(t)}{m(\underset{\sim}{x})} \le c_2 - c_4 \eta J_t \left[\frac{\alpha_{k-t}-\alpha_k}{\eta}\right]$$

Similarly

$$c_1 - c_3 \eta \min \left(J_1 \left[\frac{\alpha_1-\alpha_{s+1}}{\eta}\right], J_s \left[\frac{\alpha_s-\alpha_{s+1}}{\eta}\right]\right]$$

$$\le \Delta_t^*(s) \le c_2 - c_4 \eta J_s \left[\frac{\alpha_1-\alpha_{s+1}}{\eta}\right] \blacksquare$$

Theorem 2.5.1 : Let (a_1^*, a_2^*) denote the subset pair selected by Bayes rule and suppose that $\frac{c_1}{c_3\eta} \ge J_s(0)$, and $\frac{c_2}{c_4\eta} \ge J_t(0)$ for some $s,t \in \{1,\ldots,k-2\}$. Then $|a_1^*| \le s$, $|a_2^*| \le t$, and $|a_1^*| + |a_2^*| \le \min(s+t,k)$.

Proof : Since $J_i(z)$ is an increasing function of z and $\alpha_i - \alpha_{i+1} \le 0$ for $i = 1,\ldots,k-1$, it follows that

$$\frac{c_1}{c_3\eta} \ge J_s(0) \ge J_s \left[\frac{\alpha_s-\alpha_{s+1}}{\eta}\right]$$

and

$$\frac{c_2}{c_4\eta} \geq J_t(0) \geq J_t\left[\frac{\alpha_{k-t}-\alpha_{k-t+1}}{\eta}\right]$$

which imply that $\Delta_1(t) \geq 0$, and $\Delta_1^*(s) \geq 0$. Thus from Theorem 2.4.1, for some $m \leq t$, and $p \leq s$

$$r_1(\underset{\sim}{x},(a_1^*,a_2^*)) = r_1(\underset{\sim}{x},(a_{\{p\}}\ ,\ a^{\{m\}})),\ \text{if } m+p \leq k$$

and

$$r_1(\underset{\sim}{x},(a_1^*,a_2^*)) = \min_{k-m\leq i\leq p} r_1(\underset{\sim}{x},(a_{\{i\}},a^{\{k-i\}})),\ \text{if } m+p \geq k+1$$

which proves the required result. ■

Values of $J_i(0)$ are tabulated by Goel and Rubin (1977) for various i's.

Corollary 2.5.1 : If $\min\left[\frac{c_1}{c_3\eta}\ ,\ \frac{c_2}{c_4\eta}\right] \geq \frac{1}{\sqrt{\pi}}$, then the Bayes rule selects the subset pair ({{1}}, {{k}}) with probability one.

Corollary 2.5.2 : Suppose $\frac{c_2}{c_4\eta} \geq \frac{1}{\sqrt{\pi}}$, and $.2821 \leq \frac{c_1}{c_3\eta} < \frac{1}{\sqrt{\pi}}$, then the Bayes rule selects subset pair ({{1}}, {{k}}) if $\alpha_1 \leq \alpha_2 + \eta\, J_1^{-1}\left[\frac{c_1}{c_3\eta}\right]$, otherwise it selects subset pair ({{1}}, {2}}, {{k}}).

Proof : $\frac{c_2}{c_4\eta} \geq \frac{1}{\sqrt{\pi}} \Longrightarrow \Delta_1(1) \geq 0$.

From Table 2 of Goel and Rubin (1977), we have $J_2(0) = .2821$. Therefore

$$\frac{c_1}{c_3\eta} \geq .2821 = J_2(0) \Longrightarrow \Delta_1^*(2) \geq 0$$

Hence by Theorem 2.5.1 and Lemma 2.4.1, Bayes rule either selects subset pair ({{1}}, {{k}}) or selects pair ({{1}, {2}}, {{k}}). But subset pair ({{1}, {{k}}) is selected only if

$$\Delta_1^*(1) = c_1 - c_3\,\eta\, J_1\left[\frac{\alpha_1-\alpha_2}{\eta}\right] \geq 0,$$

that is, only if

$$\alpha_1 \leq \alpha_2 + \eta\, J_1^{-1}\left[\frac{c_1}{c_3\eta}\right]$$

and the result follows. ■

Theorem 2.5.2 : For $m, p = 1, \ldots, k-1$, suppose

(i) $\alpha_{k-m} \leq \max\left[\alpha_k + \eta\, J_1^{-1}\left[\frac{c_2}{c_4\eta}\right],\ \alpha_{k-m+1} + \eta\, J_1^{-1}\left[\frac{c_2}{c_4\eta}\right]\right]$, and

$\alpha_{k-m+1} > \alpha_k + \eta\, J_1^{-1}\left[\frac{c_2}{c_4\eta}\right]$

(ii) $\alpha_{p+1} \geq \min\left[\alpha_1 - \eta\, J_1^{-1}\left[\frac{c_1}{c_3\eta}\right],\ \alpha_p - \eta\, J_p^{-1}\left[\frac{c_1}{c_3\eta}\right]\right]$, and

$\alpha_1 > \alpha_p + \eta\, J_{p-1}^{-1}\left[\frac{c_1}{c_3\eta}\right]$.

Then Bayes rule selects subset pair (a_1^*, a_2^*) given by (2.4.2), where $J_o^{-1}(z) = -\infty$, and $\alpha_o = -\infty$.

Proof : Conditions (i) and (ii) imply that

$$c_2 - c_4\,\eta\,\min\left[J_1\left[\frac{\alpha_{k-m}-\alpha_k}{\eta}\right],\ J_m\left[\frac{\alpha_{k-m}-\alpha_{k-m+1}}{\eta}\right]\right] \geq 0,$$

$$c_2 - c_4\,\eta\, J_{m-1}\left[\frac{\alpha_{k-m+1}-\alpha_k}{\eta}\right] < 0,$$

$$c_1 - c_3\,\eta\,\min\left[J_1\left[\frac{\alpha_p-\alpha_{p+1}}{\eta}\right],\ J_p\left[\frac{\alpha_p-\alpha_{p+1}}{\eta}\right]\right] \geq 0,$$

and

$$c_1 - c_3\,\eta\, J_{p-1}\left[\frac{\alpha_1-\alpha_p}{\eta}\right]$$

and thus application of Lemma 2.5.2, gives

$$\Delta_s(m-1) < 0,\ \Delta_s(m) \geq 0\ \Delta_t^*(p-1) < 0, \text{ and } \Delta_t^*(p) \geq 0.$$

Now on using the fact that $\Delta_s(i) \le \Delta_s(i+1)$, and $\Delta_t^*(i) \le \Delta_t^*(i+1)$, we get

$$\Delta_s(i) < 0, \; i = 1,\ldots,m-1, \; \Delta_s(m) \ge 0,$$

and

$$\Delta_t^*(i) < 0, \; i = 1,\ldots,p-1, \; \Delta_t^*(p) \ge 0,$$

that is, m is the smallest integer such that $\Delta_s(.) \ge 0$, and p is the smallest integer such that $\Delta_t^*(.) \ge 0$.

Now result follow from Theorem 2.4.1. ■

Values of $-J_t^{-1}(c)$ are tabulated in Goel and Rubin (1977) for various values of t and c.

Remark: Results of this section extend to the case of multivariate normal exchangeable priors with common non-negative correlation ρ by considering $\theta_i = \theta_0'\sqrt{\rho} + \theta_i'\sqrt{(1-\rho)}$ $(i=1,\ldots,k)$, where $\theta_0', \theta_1', \ldots, \theta_k'$ are independent normal variates with common variance ξ^2 and $E(\theta_0') = 0$, $E(\theta_i') = \mu/\sqrt{(1-\rho)}$

CHAPTER - III

SIMULTANEOUS SELECTION OF EXTREME POPULATIONS : BAYES-P^* AND MINIMAX RULES

3.1 Introduction

To ensure high probability of CS, Gupta and Yang (1985) considered only those decision rules for which posterior probability of CS is at least equal to P^*, a pre-assigned number, and then derived a Bayes rule among these rules for the goal of selecting a non-empty subset of $\{\pi_1,\ldots,\pi_k\}$ containing the UEP. For selecting a non-empty subset containing the UEP, Berger (1979) derived minimax rules in the class of rules satisfying P^*-condition ($P_{\underset{\sim}{\theta}}(CS) \geq P^*$, $\forall\ \underset{\sim}{\theta} \in \Omega$) when the loss is measured by subset size. Gupta and Miescke (1986) dealt with the problem of selecting the UEP among k normal populations (with possibly unequal variances). They establish that if the risk is measured by the probability of incorrect selection then the natural decision rule which selects the population corresponding to largest observation is minimax if and only if all the variances are equal.

In this chapter we extend the results of Gupta and Yang (1985), Berger (1979), and Gupta and Miescke (1986) to the problem of simultaneous selection of the LEP and the UEP. Section 3.2 deals with the problem of finding Bayes-P^* rules (Bayes rules for which posterior probability of CS is greater than or equal to P^*, where P^* is a pre-assigned constant) with respect to a general class of loss functions. Problem of finding a minimax rule in the

class of rules satisfying P^*-condition is studied in Section 3.3. In Section 3.4, we investigate the minimaxity of the natural decision rule (which selects the populations corresponding to largest and smallest observations as the LEP and the UEP, respectively) for simultaneously selecting the lower and the upper extreme normal populations (with possibly unequal variances), when the risk is measured by the probability of incorrect selection.

3.2 Bayes-P^* Rules

In this section we continue to follow the set-up of Chapter II so that the observation X_i from π_i has a pdf $f(.;\theta_i)$ and random variables $X_1,\ldots,X_k$ are statistically independent. Consider the goal of simultaneously selecting two non-empty subsets S_L and S_U of $\{1,\ldots,k\}$ containing the indexes of the LEP and the UEP, respectively, so that our action space is given by,

$$\mathcal{A} = \{(a_1,a_2) : a_1 \subset \{1,\ldots,k\},\ a_2 \subset \{1,\ldots,k\},\ a_1 \neq \phi,\ a_2 \neq \phi\}$$

where ϕ denotes the empty set.

Let $\tau(\underset{\sim}{\theta})$ be a symmetric prior distribution on the parameter space Ω and suppose that the posterior distribution of $\underset{\sim}{\theta}$ given $\underset{\sim}{X} = \underset{\sim}{x}$ is absolutely continuous in $\underset{\sim}{\theta}$. For the problem at hand, the CS is an event which selects the index of the LEP in S_L and simultaneously selects the index of the UEP in S_U. Let $\mathcal{D}$ denotes the class of all decision rules and suppose that for a given decision rule $\delta \in \mathcal{D}$, $P(CS|\delta,\underset{\sim}{x})$ denotes the posterior probability of CS given $\underset{\sim}{X} = \underset{\sim}{x}$ using rule $\delta \in \mathcal{D}$. For a specified constant P^*, define

$$\mathcal{D}_p(P^*) = \{\delta \in \mathcal{D} : P(CS|\delta, \underset{\sim}{x}) \geq P^*, \ \forall \ \underset{\sim}{x} \in \mathcal{X}\}$$

that is, $\delta \in \mathcal{D}_p(P^*)$, if and only if

$$\sum_{\substack{i=1 \\ i \neq j}}^{k} \sum_{j=1}^{k} \psi^{\delta}_{i,j}(\underset{\sim}{x}) \ p_{ij}(\underset{\sim}{x}) \geq P^*, \quad \forall \ \underset{\sim}{x} \in \mathcal{X},$$

where

$$\psi^{\delta}_{i,j}(\underset{\sim}{x}) = \sum_{(a_1, a_2): i \in a_1, j \in a_2} \delta((a_1, a_2)|\underset{\sim}{x})$$

are the pair selection probabilities associated with decision rule δ and

$$p_{ij}(\underset{\sim}{x}) = P(\theta_i = \theta_{[1]} \text{ and } \theta_j = \theta_{[k]}|\underset{\sim}{X} = \underset{\sim}{x}).$$

Define

$$D_p(P^*) = \{\delta \in \mathcal{D}_p(P^*) : \delta \text{ is non-randomized}\}$$

Since the posterior distribution function of $\underset{\sim}{\theta}$ given $\underset{\sim}{X} = \underset{\sim}{x}$ is absolutely continuous in $\underset{\sim}{\theta}$, it is clear that

$$\sum_{\substack{i=1 \\ i \neq j}}^{k} \sum_{j=1}^{k} p_{ij}(\underset{\sim}{x}) = 1, \qquad \forall \ \underset{\sim}{x} \in \mathcal{X}$$

and hence $\mathcal{D}_p(P^*)$ is non-empty.

Remark: If $\psi^{\delta}_{i,j}$'s take on values 0 and 1 only, then

$$\delta((a_1, a_2)|\underset{\sim}{x}) = \begin{cases} 1, & \text{if } a_1 \times a_2 = \{(i,j): \psi^{\delta}_{i,j}(\underset{\sim}{x}) = 1\} \\ 0, & \text{otherwise} \end{cases}$$

where $a_1 \times a_2$ denotes the Cartesian product of the sets a_1 and a_2. Thus, a non-randomized decision rule is uniquely determined by pair selection probabilities $\psi^{\delta}_{i,j}$'s.

For example if $k=5$, $\psi^{\delta}_{1,5}(\underset{\sim}{x}) = \psi^{\delta}_{2,5}(\underset{\sim}{x}) = 1$, and all other $\psi^{\delta}_{i,j}(\underset{\sim}{x})$'s are 0, then

$$\delta((a_1,a_2)|\underset{\sim}{x}) = \begin{cases} 1, & \text{if } (a_1,a_2) = (\{1,2\},\{5\}) \\ 0, & \text{otherwise} \end{cases}$$

Definition 3.2.1

A loss function L is said to have property T_E, if

$$\left.\begin{array}{ll} \text{(i)} & L \text{ satisfies (2.2.2)}, \\ \text{(ii)} & L(\underset{\sim}{\theta},(a_1,a_2)) = L(g\underset{\sim}{\theta},(ga_1,ga_2)), \\ & \qquad \text{for all } \underset{\sim}{\theta} \in \Omega,\ (a_1,a_2) \in \mathscr{A}, \text{ and } g \in G, \\ \text{(iii)} & L(\underset{\sim}{\theta},(a_1,a_2)) \le L(\underset{\sim}{\theta},(a_1',a_2')), \text{ if } a_1 \subset a_1' \text{ and } a_2 \subset a_2' \end{array}\right\} \quad (3.2.1)$$

Our aim is to find a Bayes rule in the class $\mathbf{D_P}(P^*)$ under the above set-up. The following lemma is useful in doing so.

Lemma 3.2.1: Suppose that the pdf's $f(\cdot;\theta_i)$, $i = 1,\dots,k$ have MLR property and the prior distribution $\tau(\underset{\sim}{\theta})$ is symmetric on Ω, then

(i) $p_{ij}(\underset{\sim}{x}) \ge p_{i',j}(\underset{\sim}{x})$, for $x_{i'} \ge x_i$

(ii) $p_{ij}(\underset{\sim}{x}) \ge p_{ij'}(\underset{\sim}{x})$, for $x_j \ge x_{j'}$

Proof:

(i) $p_{ij}(\underset{\sim}{x}) - p_{i',j}(\underset{\sim}{x})$

$$= m(\underset{\sim}{x}) \int_{\Omega} (E_{ij}(\underset{\sim}{\theta}) - E_{i',j}(\underset{\sim}{\theta})) \prod_{l=1}^{k} f(x_l;\theta_l)\,d\tau(\underset{\sim}{\theta})$$

where

$$E_{ij}(\underset{\sim}{\theta}) = \begin{cases} 1, & \text{if } \theta_i = \theta_{[1]} \text{ and } \theta_j = \theta_{[k]} \\ 0, & \text{otherwise} \end{cases}$$

and $m(\underset{\sim}{x})$ is the normalizing constant.

$$p_{ij}(\underset{\sim}{x}) - p_{i',j}(\underset{\sim}{x})$$

$$= m(\underset{\sim}{x}) \Bigg[\int_{\{\underset{\sim}{\theta}:\theta_i<\theta_{i'}\}} (E_{ij}(\underset{\sim}{\theta}) - E_{i',j}(\underset{\sim}{\theta})) \prod_{e=1}^{k} f(x_e;\theta_e)d\tau(\underset{\sim}{\theta})$$

$$+ \int_{\{\underset{\sim}{\theta}:\theta_i=\theta_{i'}\}} (E_{ij}(\underset{\sim}{\theta}) - E_{i',j}(\underset{\sim}{\theta})) \prod_{e=1}^{k} f(x_e;\theta_e)d\tau(\underset{\sim}{\theta})$$

$$+ \int_{\{\underset{\sim}{\theta}:\theta_i>\theta_{i'}\}} (E_{ij}(\underset{\sim}{\theta}) - E_{i',j}(\underset{\sim}{\theta})) \prod_{e=1}^{k} f(x_e;\theta_e)d\tau(\underset{\sim}{\theta}) \Bigg]$$

Now, since τ is symmetric on Ω, and the posterior distribution of $\underset{\sim}{\theta}$ given $\underset{\sim}{X} = \underset{\sim}{x}$ is continuous, the second integral is zero and roles of θ_i and $\theta_{i'}$ can be interchanged in the third integral so that

$$p_{ij}(\underset{\sim}{x}) - p_{i',j}(\underset{\sim}{x})$$

$$= m(\underset{\sim}{x}) \int_{\{\underset{\sim}{\theta}:\theta_i<\theta_{i'}\}} (E_{ij}(\underset{\sim}{\theta}) - E_{i',j}(\underset{\sim}{\theta})) \prod_{\substack{e=1 \\ e\neq i,i'}}^{k} f(x_e;\theta_e)$$

$$(f(x_i,\theta_i)f(x_{i'};\theta_{i'})-f(x_i;\theta_{i'})f(x_{i'};\theta_i)) \; d\tau(\underset{\sim}{\theta})$$

Now using the MLR property of pdf, and noting that if $\theta_i < \theta_{i'}$, then $E_{ij}(\underset{\sim}{\theta}) \geq E_{i',j}(\underset{\sim}{\theta})$, we get

$$p_{ij}(\underset{\sim}{x}) - P_{i',j}(\underset{\sim}{x}) \geq 0.$$

(ii) Proof is similar to that (i) with obvious modifications. ■

Recall that (i) denotes the index of population associated with $X_{[i]}$ and the decision rule $\delta^{s,t}$ $(1 \le s,t \le k)$ is defined by

$$\delta^{s,t}((a_1,a_2)|\underset{\sim}{x}) = \begin{cases} 1, & \text{if } a_1=\{(1),\ldots,(s)\} \text{ and} \\ & \quad a_2=\{(k-t+1),\ldots,(k)\} \\ 0, & \text{otherwise.} \end{cases}$$

For a given P^*, $\underset{\sim}{x} \in \mathcal{X}$, let

$$\Delta_i(\underset{\sim}{x}) = \{j : \sum_{l_1=1}^{i} \sum_{l_2=k-j+1}^{k} P_{(l_1)(l_2)}(\underset{\sim}{x}) \ge P^*\},\ i = 1,\ldots,k$$

and for $\Delta_i(\underset{\sim}{x}) \ne \phi$, let $u(i)$ be the smallest integer in $\Delta_i(\underset{\sim}{x})$.

Further suppose $\mathcal{I} = \{i:\ 1 \le i \le k \text{ and } \Delta_i(\underset{\sim}{x}) \ne \phi\}$. Define

$$\mathbf{D}_o(P^*) = \{\ \delta^{i,u(i)}\ :\ i \in \mathcal{I}\ \}$$

Note: Since for any P^* and $\underset{\sim}{x} \in \mathcal{X}$, $\Delta_k(\underset{\sim}{x}) \ne \phi$, therefore $\mathbf{D}_o(P^*)$ is non-empty.

In the following theorem, we prove that $\mathbf{D}_o(P^*)$ is an essentially complete class for finding a Bayes rule in $\mathbf{D}_P(P^*)$.

Theorem 3.2.1: Suppose that the assumptions of Lemma 3.2.1 are satisfied and the loss function L has property T_E defined by (3.2.1). Then $\mathbf{D}_o(P^*)$ is essentially complete for finding a Bayes rule in $\mathbf{D}_P(P^*)$.

Proof:

Let

$$\mathbf{D}_P^*(P^*) = \{\delta \in \mathbf{D}_B\ :\ P(CS|\delta,\underset{\sim}{x}) \ge P^*\}$$

where $\mathbf{D}_B$ is the class of decision rules of the type $\delta^{s,t}$, $s,\ t = 1,\ldots,k$.

First we will show that it is enough to find a Bayes rule in $D_P^*(P^*)$. For this, let δ be any decision rule in $D_P(P^*)$ and suppose that δ selects subset pair $(\{i_1,\ldots,i_s\}, \{j_1,\ldots,j_t\})$ w.p.1. Now consider the decision rule $\delta^{s,t}$ which selects subset pair $(\{\{1\},\ldots,\{s\}\}, \{\{k-t+1\},\ldots,\{k\}\})$ w.p.1. Since $\delta \in D_P(P^*)$, we have

$$P(CS|\delta,\underset{\sim}{x}) = \sum_{l=1}^{s}\sum_{m=1}^{t} p_{i_l j_m}(\underset{\sim}{x}) \geq P^* \quad , \forall\ \underset{\sim}{x} \in \mathcal{X}$$

Now on repeatedly using (i) and (ii) of Lemma 3.2.1, we get

$$P^* \leq P(CS|\delta,\underset{\sim}{x}) \leq \sum_{l=1}^{s}\sum_{m=k-t+1}^{k} p_{\{l\}\{m\}}(\underset{\sim}{x}) = P(CS|\delta^{s,t},\underset{\sim}{x})$$

Hence $\delta^{s,t} \in D_P(P^*)$.

Now, on using Lemma 2.2.2, we have

$$\begin{aligned} r(\tau,\delta^{s,t}) &= \int_{\mathcal{X}} r_1(\underset{\sim}{x},(\{\{1\},\ldots,\{s\}\}, \{\{k-t+1\},\ldots,\{k\}\}))\, d\underset{\sim}{x} \\ &\leq \int_{\mathcal{X}} r_1(\underset{\sim}{x},(\{i_1,\ldots,i_s\}, \{j_1,\ldots,j_t\}))\, d\underset{\sim}{x} \\ &= r(\tau,\delta). \end{aligned}$$

Hence $D_P^*(P^*)$ is essentially complete for finding a Bayes rule in $D_P(P^*)$.

Now the result follows from the definition of $D_o(P^*)$ and (iii) of (3.2.1). ■

With the given conditions on the loss function no further simplification of Bayes rules is possible. As an alternative consider the class of decision rules

$$\mathcal{D}_P(P_1^*,P_2^*) = \{\delta \in \mathcal{D} : P(\pi_{(1)} \in S_L|\delta,\underset{\sim}{x}) \geq P_1^*,$$
$$P(\pi_{(k)} \in S_U|\delta,\underset{\sim}{x}) \geq P_2^*, \ \forall \ \underset{\sim}{x} \in \mathcal{X}\}$$

where P_1^* and P_2^* are pre-assigned constants. Thus, $\delta \in \mathcal{D}_P(P_1^*,P_2^*)$, if and only if

$$\sum_{i=1}^{k} \psi_{i.}^{\delta}(\underset{\sim}{x})\ p_{i.}(\underset{\sim}{x}) \geq P^*, \quad \forall \ \underset{\sim}{x} \in \mathcal{X}$$

and

$$\sum_{j=1}^{k} \psi_{.j}^{\delta}(\underset{\sim}{x})\ p_{.j}(\underset{\sim}{x}) \geq P_2^*, \quad \forall \ \underset{\sim}{x} \in \mathcal{X},$$

where

$$\psi_{i.}^{\delta}(\underset{\sim}{x}) = \sum_{(a_1,a_2):i\in a_1} \delta((a_1,a_2)|\underset{\sim}{x}),$$

$$\psi_{.j}^{\delta}(\underset{\sim}{x}) = \sum_{(a_1,a_2):j\in a_2} \delta((a_1,a_2)|\underset{\sim}{x}),$$

$$p_{i.}(\underset{\sim}{x}) = P(\theta_i = \theta_{[1]}|\underset{\sim}{X} = \underset{\sim}{x}),$$

and

$$p_{.j}(\underset{\sim}{x}) = P(\theta_j = \theta_{[k]}|\underset{\sim}{X} = \underset{\sim}{x}).$$

Let

$$\mathbf{D}_{\mathbf{P}}(P_1^*,P_2^*) = \{\delta \in \mathcal{D}_P(P_1^*,P_2^*) : \delta \text{ is non-randomized}\}.$$

Since the posterior distribution function of $\underset{\sim}{\theta}$ given $\underset{\sim}{X} = \underset{\sim}{x}$ is absolutely continuous in $\underset{\sim}{\theta}$, it is clear that

$$\sum_{i=1}^{k} p_{i.}(\underset{\sim}{x}) = 1, \text{ and } \sum_{j=1}^{k} p_{.j}(\underset{\sim}{x}) = 1$$

and hence $\mathbf{D}_{\mathbf{P}}(P_1^*,P_2^*)$ is non-empty.

Remark: For a fixed P^*, there exist P_1^*, P_2^* such that $\mathbf{D_P}(P_1^*,P_2^*) \subset \mathbf{D_P}(P^*)$ ($P_1^* = P_2^* = \frac{1+P^*}{2}$ is one such possible choice).

Now we will find a Bayes rule in the restricted class $\mathbf{D_P}(P_1^*,P_2^*)$. To obtain a Bayes rule in $\mathbf{D_P}(P_1^*,P_2^*)$, we need the following lemma.

Lemma 3.2.2: Suppose that the pdf's $f(.;\theta_i)$, $i = 1,\ldots,k$ have MLR property and the prior distribution $\tau(\underset{\sim}{\theta})$ is symmetric on Ω. Then

(i) $p_{i.}(\underset{\sim}{x}) \geq p_{i'.}(\underset{\sim}{x})$, for $x_{i'} \geq x_i$

(ii) $p_{.j}(\underset{\sim}{x}) \geq p_{.j'}(\underset{\sim}{x})$, for $x_j \geq x_{j'}$

Proof: Since

$$p_{i.}(\underset{\sim}{x}) = \sum_{j=1}^{k} p_{ij}(\underset{\sim}{x}) \quad \text{and} \quad p_{.j}(\underset{\sim}{x}) = \sum_{i=1}^{k} p_{ij}(\underset{\sim}{x})$$

the results follow from Lemma 3.2.1. ∎

For given $\underset{\sim}{x} \in \mathcal{X}$, let m, $p \in \{1,\ldots,k\}$ be such that

$$\left.\begin{aligned} m &= \min\{i : \sum_{l_1=1}^{i} p_{\{l_1\}.}(\underset{\sim}{x}) \geq P_1^*\} \\ p &= \min\{j : \sum_{l_1=k-j+1}^{k} p_{.\{l_1\}}(\underset{\sim}{x}) \geq P_2^*\} \end{aligned}\right\} \tag{3.2.2}$$

In the following theorem, we prove that the decision rule $\delta^{m,p}$ is Bayes in $\mathbf{D_P}(P_1^*,P_2^*)$ where m and p are defined by (3.2.2).

Theorem 3.2.2: Suppose that the assumptions of Lemma 3.2.1 are satisfied and the loss function L has property T_E defined by (3.2.1). Let $m,p \in \{1,\dots,k\}$ be defined by (3.2.2). Then the decision rule $\delta^{m,p}$ is Bayes in $\mathbf{D_P}(P_1^*,P_2^*)$.

Proof: Let

$$\mathbf{D_P^*}(P_1^*,P_2^*) = \{\delta \in \mathbf{D_B} : P(\pi_{(1)} \in S_L|\delta,\underset{\sim}{x}) \geq P_1^*, \text{ and}$$
$$P(\pi_{(k)} \in S_U|\delta,\underset{\sim}{x}) \geq P_2^*, \ \forall \ \underset{\sim}{x} \in \mathcal{X}\}$$

where $\mathbf{D_B}$ is the class of decision rules of the type $\delta^{s,t}$, s, t = 1,...,k.

First we will show that it is enough to find a Bayes rule in $\mathbf{D_P^*}(P_1^*,P_2^*)$. For this, let δ be any decision rule in $\mathbf{D_P}(P_1^*,P_2^*)$ and suppose that δ selects subset pair $(\{i_1,\dots,i_s\}, \{j_1,\dots,j_t\})$ w.p.1. Now consider the decision rule $\delta^{s,t}$ which selects subset pair $(\{\{1\},\dots,\{s\}\}, \{\{k-t+1\},\dots,\{k\}\})$ w.p.1. Since $\delta \in \mathbf{D_P}(P_1^*,P_2^*)$, we have

$$\sum_{l=1}^{s} p_{i_l \cdot}(\underset{\sim}{x}) \geq P_1^* \quad , \forall \ \underset{\sim}{x} \in \mathcal{X}$$

and

$$\sum_{l=1}^{t} p_{\cdot j_l}(\underset{\sim}{x}) \geq P_2^* \quad , \forall \ \underset{\sim}{x} \in \mathcal{X}$$

Now on repeatedly using (i) and (ii) of Lemma 3.2.1, we get

$$P_1^* \leq \sum_{l=1}^{s} p_{i_l \cdot}(\underset{\sim}{x}) \leq \sum_{l=1}^{s} p_{\{l\} \cdot}(\underset{\sim}{x})$$

and

$$P_2^* \leq \sum_{l=1}^{t} p_{\cdot j_l}(\underset{\sim}{x}) \leq \sum_{l=k-t+1}^{k} p_{\cdot \{l\}}(\underset{\sim}{x})$$

Thus, $\delta^{s,t} \in \mathbf{D_P}(P_1^*, P_2^*)$

Now, on using Lemma 2.2.2, we have

$$r(\tau, \delta^{s,t}) = \int_{\mathcal{X}} r_1(\underset{\sim}{x}, (\{\{1\}, \ldots, \{s\}\}, \{\{k-t+1\}, \ldots, \{k\}\})) \, d\underset{\sim}{x}$$

$$\leq \int_{\mathcal{X}} r_1(\underset{\sim}{x}, (\{i_1, \ldots, i_s\}, \{j_1, \ldots, j_t\})) \, d\underset{\sim}{x}$$

$$= r(\tau, \delta).$$

Hence $\mathbf{D_P^*}(P_1^*, P_2^*)$ is essentially complete for finding a Bayes rule in $\mathbf{D_P}(P_1^*, P_2^*)$.

Now, since among all rules in $\mathbf{D_P^*}(P_1^*, P_2^*)$, $\delta^{m,p}$ selects the smallest sized S_L and S_U, $\delta^{m,p}$ is Bayes in $\mathbf{D_P}(P_1^*, P_2^*)$. ■

Let $E(|S_L| \, |S_U| \; | \delta, \underset{\sim}{x})$ denotes the posterior expectation of product of sizes of subsets selected by $\delta \in \mathbf{D_P}(P_1^*, P_2^*)$. For any decision rule $\delta \in \mathbf{D_P}(P_1^*, P_2^*)$, the posterior efficiency of δ, given observation $\underset{\sim}{X} = \underset{\sim}{x}$, is defined by

$$\mathrm{eff}(\delta | \underset{\sim}{x}) = \frac{P(CS | \delta, \underset{\sim}{x})}{E(|S_L| \, |S_U| \; | \delta, \underset{\sim}{x})} .$$

The expectation of $\mathrm{eff}(\delta | \underset{\sim}{X})$ is the efficiency of the decision rule δ and is denoted by $\mathrm{eff}(\delta)$. A decision rule $\delta \in \mathbf{D_P}(P_1^*, P_2^*)$ is called most efficient in $\mathbf{D_P}(P_1^*, P_2^*)$ if $\mathrm{eff}(\delta) \geq \mathrm{eff}(\delta')$ for all $\delta' \in \mathbf{D_P}(P_1^*, P_2^*)$.

The following theorem shows that decision rule $\delta^{m,p}$ is most efficient in $\mathbf{D_P}(P_1^*, P_2^*)$.

<u>**Theorem 3.2.3:**</u> Suppose that assumptions of Theorem 3.2.2 hold, then the decision rule $\delta^{m,p}$ is most efficient in $\mathbf{D_P}(P_1^*, P_2^*)$.

Proof: Since, among all decision rules in $D_P(P_1^*, P_2^*)$, given any observation $\underset{\sim}{X} = \underset{\sim}{x}$, $\delta^{m,p}$ has minimum sizes of the selected subsets S_L and S_U, any other decision rule δ' in $D_P(P_1^*, P_2^*)$ should have its sizes of S_L and S_U equal to $m+c_1$ and $p+c_2$, respectively, $c_1, c_2 \geq 0$. Now

$$(m+c_1)(p+c_2)\mathrm{eff}(\delta' \mid \underset{\sim}{x}) = \sum_{l_1=1}^{m+c_1} \sum_{l_2=k-p-c_2+1}^{k} P_{\{l_1\}\{l_2\}}(\underset{\sim}{x})$$

$$= \Bigg[\sum_{l_1=1}^{m} \sum_{l_2=k-p-c_2+1}^{k-p} P_{\{l_1\}\{l_2\}}(\underset{\sim}{x})$$

$$+ \sum_{l_1=1}^{m} \sum_{l_2=k-p+1}^{k} P_{\{l_1\}\{l_2\}}(\underset{\sim}{x})$$

$$+ \sum_{l_1=m+1}^{m+c_1} \sum_{l_2=k-p-c_2+1}^{k-p} P_{\{l_1\}\{l_2\}}(\underset{\sim}{x})$$

$$+ \sum_{l_1=m+1}^{m+c_1} \sum_{l_2=k-p+1}^{k} P_{\{l_1\}\{l_2\}}(\underset{\sim}{x})\Bigg]$$

$$\leq \Bigg[c_2 \sum_{l_1=1}^{m} P_{\{l_1\}\{k-p\}}(\underset{\sim}{x})$$

$$+ \sum_{l_1=1}^{m} \sum_{l_2=k-p+1}^{k} P_{\{l_1\}\{l_2\}}(\underset{\sim}{x})$$

$$+ c_1 c_2 P_{\{m+1\}\{k-p\}}(\underset{\sim}{x})$$

$$+ c_1 \sum_{l_2=k-p+1}^{k} P_{\{m+1\}\{l_2\}}(\underset{\sim}{x}) \Bigg]$$

(using Lemma 3.2.1)

$$\le \left[\left(\frac{c_2}{p} + 1 + \frac{c_1 c_2}{mp} + \frac{c_1}{m}\right) \sum_{l_1=1}^{m} \sum_{l_2=k-p+1}^{k} p_{\{l_1\}\{l_2\}}(\underset{\sim}{x})\right]$$

$$= \frac{(m+c_1)(p+c_2)}{mp} \sum_{l_1=1}^{m} \sum_{l_2=k-p+1}^{k} p_{\{l_1\}\{l_2\}}(\underset{\sim}{x})$$

$$= (m + c_1)\ (p + c_2)\ \mathrm{eff}\ (\delta^{m,p}|\underset{\sim}{x}).$$

Thus,

$$\mathrm{eff}(\delta'\,|\underset{\sim}{x}) \le \mathrm{eff}(\delta^{m,p}|\underset{\sim}{x}) \qquad \forall\ \underset{\sim}{x} \in \mathcal{X},$$

and therefore

$$\mathrm{eff}(\delta') = E(\mathrm{eff}(\delta'\,|\underset{\sim}{X})) \le E(\mathrm{eff}(\delta^{m,p}|\underset{\sim}{X})) = \mathrm{eff}(\delta^{m,p}) \qquad \blacksquare$$

3.3 Minimax Rules : Specific Loss Function

In this section we prove that the selection procedure of Dudewicz and Mishra (1987) is minimax with respect to a simple loss function.

Suppose random vector $\underset{\sim}{X} = (X_1,\dots,X_k)$ has joint density $g_{\underset{\sim}{X}}(\underset{\sim}{x},\underset{\sim}{\theta})$ and we are interested in selecting a subset pair (S_L,S_U) such that S_L contains the LEP and S_U contains the UEP. Consider the loss function given by

$$L(\underset{\sim}{\theta},(a_1,a_2)) = |a_1|\ |a_2| - |a_1 \cap a_2| \tag{3.3.1}$$

For a given P^* $\left(\frac{1}{k(k-1)} < P^* < 1\right)$, define

$$\mathcal{D}(P^*) = \{\delta \in \mathcal{D} : P_{\underset{\sim}{\theta}}(CS|\delta) \ge P^*,\ \forall\ \underset{\sim}{\theta} \in \Omega\}$$

where as before, CS denotes the event $\{\Pi_{(1)} \in S_L$ and $\Pi_{(k)} \in S_U\}$.

Suppose that

$$\Omega_{ij} = \{\underset{\sim}{\theta} \in \Omega : \theta_i = \theta_{[1]} \text{ and } \theta_j = \theta_{[k]}\},\ i,j = 1,\ldots,k.$$

If a parameter $\underset{\sim}{\theta}$ point could be placed in more than one Ω_{ij}, then the point is arbitrarily put in one of the sets. This is done to ensure that $\{\Omega_{ij}\}$ forms a partition of Ω. Suppose

$$\Omega_E = \{\underset{\sim}{\theta} \in \Omega:\ \underset{\sim}{\theta} \in \overline{\Omega}_{ij},\ \text{for all } i,j\}$$

where $\overline{\Omega}_{ij}$ denotes the closure of the set Ω_{ij}. For $\delta \in \mathcal{D}$, let

$$R(\underset{\sim}{\theta},\delta) = E_{\underset{\sim}{\theta}}\left[\sum_{(a_1,a_2)\in\mathcal{A}} L(\underset{\sim}{\theta},(a_1,a_2))\ \delta((a_1,a_2)|\underset{\sim}{X})\right]$$

denote the risk associated with the decision rule δ. We desire to find a rule δ^*, which is minimax in $\mathcal{D}(P^*)$, i.e. a rule δ^* which satisfies

$$\sup_{\Omega} R(\underset{\sim}{\theta},\delta^*) = \inf_{\delta\in\mathcal{D}(P^*)} \sup_{\Omega} R(\underset{\sim}{\theta},\delta).$$

To find such a δ^*, we will first find the minimax value of the problem. Following theorem gives the minimax value of the problem.

Theorem 3.3.1: Suppose Ω_E is non-empty and $P_{\underset{\sim}{\theta}}$ (Select π_i in S_L and π_j in $S_U|\delta$) is continuous at $\underset{\sim}{\theta}_E \in \Omega_E$, for all $\delta \in \mathcal{D}$ and for all $i \neq j$. Then the minimax value with respect to the loss function given by (3.3.1) is $k(k-1)P^*$.

Proof: Recall that

$$\psi^{\delta}_{i,j}(\underset{\sim}{x}) = \sum_{(a_1,a_2):i\in a_1,\,j\in a_2} \delta((a_1,a_2)|\underset{\sim}{x})$$

Then

$$R(\underset{\sim}{\theta},\delta) = E_{\underset{\sim}{\theta}}\Big[\sum_{(a_1,a_2)\in\mathcal{A}} (|a_1|\ |a_2| - |a_1 \cap a_2|)\ \delta((a_1,a_2)|\underset{\sim}{X})\Big]$$

$$= E_{\underset{\sim}{\theta}}\Big[\sum_{(a_1,a_2)\in\mathcal{A}} \ \sum_{(i,j):i\in a_1,\, j\in a_2} \delta((a_1,a_2|\underset{\sim}{X})$$

$$- \sum_{(a_1,a_2)\in\mathcal{A}} \ \sum_{i:i\in a_1\cap a_2} \delta((a_1,a_2)|\underset{\sim}{X})\Big]$$

$$= E_{\underset{\sim}{\theta}}\Big[\sum_{i=1}^{k}\sum_{j=1}^{k} \ \sum_{(a_1,a_2):i\in a_1,\, j\in a_2} \delta((a_1,a_2)|\underset{\sim}{X})$$

$$-\sum_{i=1}^{k} \ \sum_{(a_1,a_2):i\in a_1\cap a_2} \delta((a_1,a_2)|\underset{\sim}{X})\Big]$$

$$= E_{\underset{\sim}{\theta}}\ \Big[\sum_{i=1}^{k}\sum_{j=1}^{k} \psi^{\delta}_{i,j}(\underset{\sim}{X}) - \sum_{i=1}^{k} \psi^{\delta}_{i,i}(\underset{\sim}{X})\Big]$$

Therefore,

$$R(\underset{\sim}{\theta},\delta) = E_{\underset{\sim}{\theta}}\Big[\sum_{i=1}^{k}\sum_{\substack{j=1\\ i\neq j}}^{k} \psi^{\delta}_{i,j}(\underset{\sim}{X})\Big] \tag{3.3.2}$$

Consider the no-data rule δ_1, defined by

$$\delta_1((a_1,a_2)|\underset{\sim}{x}) = \begin{cases} \dfrac{1-P^*}{k(k-1)-1} & ,\ a_1=\{i\},\ a_2=\{j\},\ i,j=1,\dots k,\ i\neq j \\[2ex] \dfrac{k(k-1)P^*-1}{k(k-1)-1} & ,\ a_1=a_2=\{1,\dots,k\}\ ,\quad \forall\ \underset{\sim}{x}\in\mathcal{X} \\[2ex] 0 & ,\ \text{otherwise.} \end{cases}$$

Clearly δ_1 is a valid decision rule and

$$\psi^{\delta_1}_{i,j}(\underset{\sim}{x}) = P^*,\ \forall\ \underset{\sim}{x}\in\mathcal{X},\ i,j = 1,\dots,k,\ i\neq j.$$

Therefore,

$$R(\underset{\sim}{\theta},\delta_1) = k(k-1)P^*\ ,\quad \forall\ \underset{\sim}{\theta}\in\Omega$$

and the minimax value

$$\inf_{\delta\in\mathcal{D}(P^*)} \sup_{\underset{\sim}{\theta}} R(\underset{\sim}{\theta},\delta) \le \sup_{\underset{\sim}{\theta}} R(\underset{\sim}{\theta},\delta_1) = k(k-1)P^*.$$

Let $\{\underset{\sim}{\theta}^n_{ij},\ n = 1,2,\ldots\}$ be a sequence in Ω_{ij} which converges to $\underset{\sim}{\theta}_E$ ($\underset{\sim}{\theta}_E \in \bar{\Omega}_{ij}$ guarantees the existence of such a sequence). Since $\underset{\sim}{\theta}^n_{ij} \in \Omega_{ij}$, selection of π_i in S_L and π_j in S_U is a correct selection at $\underset{\sim}{\theta}^n_{ij}$. Suppose $\delta \in \mathcal{D}(P^*)$, then using the continuity of $P_{\underset{\sim}{\theta}}(\text{Select } \pi_i \text{ in } S_L \text{ and } \pi_j \text{ in } S_U|\delta)$ at $\underset{\sim}{\theta}_E$, we have for all $i \ne j$,

$$\begin{aligned} E_{\underset{\sim}{\theta}_E}(\psi^{\delta}_{i,j}(\underset{\sim}{X})) &= P_{\underset{\sim}{\theta}_E}(\text{Select } \pi_i \text{ in } S_L \text{ and } \pi_j \text{ in } S_U|\delta) \\ &= \lim_{n\to\infty} P_{\underset{\sim}{\theta}^n_{ij}}(\text{Select } \pi_i \text{ in } S_L \text{ and } \pi_j \text{ in } S_U|\delta) \\ &= \lim_{n\to\infty} P_{\underset{\sim}{\theta}^n_{ij}}(CS|\delta) \ge P^* \end{aligned} \tag{3.3.3}$$

This is true for all $\delta \in \mathcal{D}(P^*)$ and for all $i \ne j$. Hence for any $\delta \in \mathcal{D}(P^*)$

$$\begin{aligned} \sup_{\Omega} R(\underset{\sim}{\theta},\delta) &\ge R(\underset{\sim}{\theta}_E,\delta) \\ &= E_{\underset{\sim}{\theta}_E}\Big[\sum_{\substack{i=1}}^{k} \sum_{\substack{j=1\\ i\ne j}}^{k} \psi^{\delta}_{i,j}(\underset{\sim}{X})\Big] \\ &\ge k(k-1)P^* \qquad \text{(from (3.3.3))} \end{aligned}$$

which implies that minimax value is not less than $k(k-1)P^*$ and the result follows. ∎

Following theorem due to Berger (1977) provides conditions under which the continuity assumption of Theorem 3.3.1 is satisfied.

Theorem 3.3.2:(Berger). Suppose $\{g_{\underset{\sim}{X}}(\underset{\sim}{x};\underset{\sim}{\theta}),\ \underset{\sim}{\theta} \in \Omega\}$ are densities with respect to a measure μ which satisfy

(i) as $\underset{\sim}{\theta} \longrightarrow \underset{\sim}{\theta}_E$, $g_{\underset{\sim}{X}}(\underset{\sim}{x};\underset{\sim}{\theta}) \longrightarrow g_{\underset{\sim}{X}}(\underset{\sim}{x};\underset{\sim}{\theta}_E)$, a.e. μ,

or (ii) as $\underset{\sim}{\theta} \longrightarrow \underset{\sim}{\theta}_E$, $g_{\underset{\sim}{X}}(\underset{\sim}{x};\underset{\sim}{\theta}) \longrightarrow g_{\underset{\sim}{X}}(\underset{\sim}{x};\underset{\sim}{\theta}_E)$ in μ measure.

Then for any bounded measurable function $\psi(\underset{\sim}{x})$, $E_{\underset{\sim}{\theta}}(\psi(\underset{\sim}{X}))$ is continuous at $\underset{\sim}{\theta}_E$.

Example 3.3.1: Suppose $\underset{\sim}{\theta}$ is a location (scale) parameter and $g_{\underset{\sim}{X}}(\underset{\sim}{x};\underset{\sim}{\theta}) = g(\underset{\sim}{x}-\underset{\sim}{\theta})$ $\left[g_{\underset{\sim}{X}}(\underset{\sim}{x};\underset{\sim}{\theta}) = \frac{1}{\theta_1} \cdots \frac{1}{\theta_k}\, g\left(\frac{x_1}{\theta_1}, \frac{x_2}{\theta_2}, \ldots, \frac{x_k}{\theta_k}\right)\right]$ is the pdf of $\underset{\sim}{X}$ with respect to Lebesgue measure, μ, on $\mathbb{R}^k$ $((0,\infty)\times \ldots \times(0,\infty))$. Suppose $g(\underset{\sim}{x})$ is continuous a.e. μ. Let A be the set of discontinuities of g. Then for fixed $\underset{\sim}{\theta}_E \in \mathbb{R}^k$ $((0,\infty)\times \ldots \times(0,\infty))$, the set of $\underset{\sim}{x}$ for which $g(\underset{\sim}{x}-\underset{\sim}{\theta})\left[\frac{1}{\theta_1}, \ldots, \frac{1}{\theta_k}\, g\left(\frac{x_1}{\theta_1}, \ldots, \frac{x_k}{\theta_k}\right)\right]$ is not continuous at $\underset{\sim}{\theta}_E$ is $\{\underset{\sim}{x}:\underset{\sim}{x} = \underset{\sim}{z} + \underset{\sim}{\theta}_E,\ \underset{\sim}{z} \in A\} = A + \underset{\sim}{\theta}_E$ $(\{\underset{\sim}{x}:x_i = z_i\theta_i,\ i = 1,\ldots,k, \underset{\sim}{z} \in A\} = A.\underset{\sim}{\theta}_E)$ and $\mu(A) = 0$ implies that $\mu(A + \underset{\sim}{\theta}_E) = \mu(A) = 0$ $(\mu(A.\underset{\sim}{\theta}_E) = \mu(A) = 0)$. So condition (i) of Theorem 3.3.2 is satisfied for every $\underset{\sim}{\theta}_E \in \mathbb{R}^k$ $((0,\infty) \times \ldots \times (0,\infty))$ and $P_{\underset{\sim}{\theta}}$(Select π_i in S_L and π_j in $S_U|\delta$) is a continuous function of $\underset{\sim}{\theta}$ on $\mathbb{R}^k$ $((0,\infty)\times \ldots \times(0,\infty))$ for any δ.

Example 3.3.2: Suppose $\underset{\sim}{X}$ has a multinomial distribution with cell probabilities $\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_k)$. The sample space is $\mathcal{X} = \{(x_1, \ldots, x_k) : x_i \in \{0, 1, \ldots, N\}, \sum_{i=1}^{k} x_i = N\}$ and the parameter space is $\Omega = \{\underset{\sim}{\theta} : \theta_i \geq 0, i=1, \ldots, k \text{ and } \sum_{i=1}^{k} \theta_i = 1\}$. The density with respect to counting measure on $\mathcal{X}$ is given by

$$g_{\underset{\sim}{X}}(\underset{\sim}{x}; \underset{\sim}{\theta}) = \frac{\lfloor N}{\lfloor x_1 \cdots \lfloor x_k} \theta_1^{x_1} \cdots \theta_k^{x_k}, \qquad \underset{\sim}{x} \in \mathcal{X}.$$

For every $\underset{\sim}{x} \in \mathcal{X}$ this is a polynomial in $\underset{\sim}{\theta}$ so is continuous in $\underset{\sim}{\theta}$. So again, condition (i) of Theorem 3.3.2 is satisfied.

Now suppose that $X_1, \ldots, X_k$ are $k(\geq 2)$ independent r.v.'s and suppose that the observation X_i from π_i has a pdf $f_{X_i}(x_i; \theta_i) = f(x_i - \theta_i)$, $i = 1, \ldots, k$. Mishra and Dudewicz (1987) propose following decision rule R_{MD}.

R_{MD} : Select π_i in S_L if and only if $X_i \leq X_{[1]} + d_1$

and

Select π_j in S_U if and only if $X_j \geq X_{[k]} - d_1$

where d_1 is chosen such that

$$\inf_{\Omega} P_{\underset{\sim}{\theta}}(CS|R_{MD}) = \int_{-\infty}^{\infty} \int_{-\infty}^{s+d_1} \left[F(s+d_1) - F(t-d_1)\right]^{k-2} f(s)\, f(t)\, dt\, ds$$

Here F denotes the distribution function corresponding to f.

Lemma 3.3.1: If $f_{X_i}(x_i;\theta_i) = f(x_i-\theta_i)$, $i = 1,\dots,k$ and pdf's $f_{X_i}(x_i;\theta_i)$ have MLR property, then supremum of $R(\underset{\sim}{\theta},R_{MD})$ is attained when $\theta_1 = \dots = \theta_k$ and $\sup_{\Omega} R(\underset{\sim}{\theta},R_{MD}) = k(k-1)P^*$.

Proof:
$$R(\underset{\sim}{\theta},R_{MD}) = \sum_{i=1}^{k}\sum_{\substack{j=1\\ i\neq j}}^{k} P_{\underset{\sim}{\theta}}(\text{Select } \pi_i \text{ in } S_L \text{ and } \pi_j \text{ in } S_U | R_{MD})$$

$$= \sum_{i=1}^{k}\sum_{\substack{j=1\\ i\neq j}}^{k} P_{\underset{\sim}{\theta}}(\text{Select } \pi_{(i)} \text{ in } S_L \text{ and } \pi_{(j)} \text{ in } S_U | R_{MD})$$

$$= \sum_{i=1}^{k}\sum_{\substack{j=1\\ i\neq j}}^{k} P_{\underset{\sim}{\theta}}(X_{(i)} \le X_{[1]}+d_1,\ X_{(j)} \ge X_{[k]}-d_1)$$

$$= \sum_{i=1}^{k}\sum_{\substack{j=1\\ i\neq j}}^{k} \int_{-\infty}^{\infty} \int_{-\infty}^{s+\theta_{[j]}-\theta_{[i]}+d_1} \prod_{\substack{l=1\\ l\neq i,j}}^{k} (F(s+\theta_{[j]}-\theta_{[l]}+d_1)$$
$$- F(t+\theta_{[i]} - \theta_{[l]} - d_1))\ f(s)\ f(t)\ dt\ ds \tag{3.3.4}$$

For $k > m \ge 1$, and for fixed $\theta_{[m+1]}, \dots, \theta_{[k]}$, let

$$Q_1(\theta) = E_{\underset{\sim}{\theta}}\Big[\sum_{(a_1,a_2)\in\mathscr{A}} L(\underset{\sim}{\theta},(a_1,a_2))\delta((a_1,a_2)|\underset{\sim}{X})|\theta_{[1]}=\dots=\theta_{[m]}=\theta,R_{MD}\Big]$$

$$= \sum_{i=1}^{m}\sum_{\substack{j=1\\ i\neq j}}^{m} \int_{-\infty}^{\infty}\int_{-\infty}^{s+d_1} \prod_{\substack{l=1\\ l\neq i,j}}^{k} (F(s+\theta-\theta_{[l]}+d_1)-F(t+\theta-\theta_{[l]}-d_1))$$
$$f(s)\ f(t)\ dt\ ds$$

$$+ \sum_{i=1}^{m} \sum_{j=m+1}^{k} \int_{-\infty}^{\infty} \int_{-\infty}^{s+\theta_{[j]}-\theta+d_1} \prod_{\substack{l=1 \\ l\neq i,j}}^{k} (F(s+\theta_{[j]}-\theta_{[l]}+d_1)$$

$$- F(t+\theta-\theta_{[l]}-d_1))f(s)f(t)dt\ ds$$

$$+ \sum_{i=m+1}^{k} \sum_{j=1}^{m} \int_{-\infty}^{\infty} \int_{-\infty}^{s+\theta-\theta_{[i]}+d_1} \prod_{\substack{l=1 \\ l\neq i,j}}^{k} (F(s+\theta-\theta_{[l]}+d_1)$$

$$- F(t+\theta_{[i]}-\theta_{[l]}-d_1))f(s)f(t)dt\ ds$$

$$+ \sum_{\substack{i=m+1 \\ i\neq j}}^{k} \sum_{j=m+1}^{k} \int_{-\infty}^{\infty} \int_{-\infty}^{s+\theta_{[j]}-\theta_{[i]}+d_1} \prod_{\substack{l=1 \\ l\neq i,j}}^{k} (F(s+\theta_{[j]}-\theta_{[l]}+d_1)$$

$$- F(t+\theta_{[i]}-\theta_{[l]}-d_1))f(s)f(t)dt\ ds$$

$$= m(m-1) \int_{-\infty}^{\infty} \int_{-\infty}^{s+d_1} (F(s+d_1)-F(t-d_1))^{m-2} \prod_{l=m+1}^{k} (F(s+\theta-\theta_{[l]}+d_1)$$

$$- F(t+\theta-\theta_{[l]}-d_1))\ f(s)f(t)dt\ ds$$

$$+ m \sum_{j=m+1}^{k} \int_{-\infty}^{\infty} \int_{-\infty}^{s+\theta_{[j]}-\theta+d_1} (F(s+\theta_{[j]}-\theta+d_1) - F(t-d_1))^{m-1}$$

$$\prod_{\substack{l=m+1 \\ l\neq j}}^{k} (F(s+\theta_{[j]}-\theta_{[l]}+d_1) - F(t+\theta-\theta_{[l]}-d_1))f(s)f(t)dt\ ds$$

$$+ m \sum_{i=m+1}^{k} \int_{-\infty}^{\infty} \int_{-\infty}^{s+\theta-\theta_{[i]}+d_1} (F(s+d_1) - F(t+\theta_{[i]}-\theta-d_1))^{m-1}$$

$$\prod_{\substack{l=m+1 \\ l\neq i}}^{k} (F(s+\theta-\theta_{[l]}+d_1) - F(t+\theta_{[i]}-\theta_{[l]}-d_1))f(s)f(t)dt\ ds$$

$$+ \sum_{\substack{i=m+1 \\ i\neq j}}^{k} \sum_{j=m+1}^{k} \int_{-\infty}^{\infty} \int_{-\infty}^{s+\theta_{[j]}-\theta_{[i]}+d_1} (F(s+\theta_{[j]}-\theta+d_1) - F(t+\theta_{[i]}-\theta-d_1))^m$$

$$\prod_{\substack{l=m+1 \\ l\neq i,j}}^{k} (F(s+\theta_{[j]}-\theta_{[l]}+d_1) - F(t+\theta_{[i]}-\theta_{[l]}-d_1))f(s)f(t)dt\ ds$$

We now show that $Q_1(\theta)$ is a non-decreasing function of θ which proves that $Q_1(\theta)$ is maximum where $\theta = \theta_{[m+1]}$ and since this holds for any $m < k$, the first assertion of the theorem will follow. To show that $Q_1(\theta)$ is a non-decreasing we differentiate $Q_1(\theta)$ with respect to θ and show that the result is positive for $\theta < \theta_{[m+1]}$.

$$\frac{d}{d\theta} Q_1(\theta)$$

$$= m(m-1) \int_{-\infty}^{\infty} \int_{-\infty}^{s+d_1} (F(s+d_1)-F(t-d_1))^{m-2} \sum_{i=m+1}^{k} \prod_{\substack{l=m+1 \\ l\neq i}}^{k} (F(s+\theta-\theta_{[l]}+d_1)$$

$$-F(t+\theta-\theta_{[l]}-d_1)) (f(s+\theta-\theta_{[i]}+d_1)-f(t+\theta-\theta_{[i]}-d_1)) f(s)f(t)dt\ ds$$

$$+m \sum_{\substack{i=m+1 \\ i\neq j}}^{k} \sum_{j=m+1}^{k} \int_{-\infty}^{\infty} \int_{-\infty}^{s+\theta_{[j]}-\theta_{[i]}+d_1} (F(s+\theta_{[j]}-\theta+d_1) - F(t+\theta_{[i]}-\theta-d_1))^{m-1}$$

$$\prod_{\substack{l=m+1 \\ l\neq i,j}}^{k} (F(s+\theta_{[j]}-\theta_{[l]}+d_1) - F(t+\theta_{[i]}-\theta_{[l]}-d_1))$$

$$(-f(s+\theta_{[j]}-\theta+d_1)-f(t+\theta_{[i]}-\theta-d_1))\ f(s)f(t)dt\ ds$$

$$+ \; m(m-1) \sum_{j=m+1}^{k} \int_{-\infty}^{\infty} \int_{-\infty}^{s+\theta_{[j]}-\theta+d_1} (F(s+\theta_{[j]}-\theta+d_1) - F(t-d_1))^{m-2}$$

$$\prod_{\substack{l=m+1 \\ l \neq j}}^{k} (F(s+\theta_{[j]}-\theta_{[l]}+d_1) - F(t+\theta-\theta_{[l]}-d_1))$$

$$(-f(s+\theta_{[j]}-\theta+d_1)) \; f(s)f(t)dt\; ds$$

$$+ \; m \sum_{j=m+1}^{k} \int_{-\infty}^{\infty} \int_{-\infty}^{s+\theta_{[j]}-\theta+d_1} (F(s+\theta_{[j]}-\theta+d_1) - F(t-d_1))^{m-1}$$

$$\sum_{\substack{i=m+1 \\ i \neq j}}^{k} \prod_{\substack{l=m+1 \\ l \neq i,j}}^{k} (F(s+\theta_{[j]}-\theta_{[l]}+d_1) - F(t+\theta-\theta_{[l]}-d_1))$$

$$(-f(t+\theta-\theta_{[i]}-d_1)) \; f(s)f(t)dt\; ds$$

$$- \; m \sum_{j=m+1}^{k} \int_{-\infty}^{\infty} (F(s+\theta_{[j]}-\theta+d_1) - F(s+\theta_{[j]}-\theta))^{m-1}$$

$$\prod_{\substack{l=m+1 \\ l \neq j}}^{k} (F(s+\theta_{[j]}-\theta_{[l]}+d_1) - F(s+\theta_{[j]}-\theta_{[l]}))$$

$$\times \; f(s)f(s+\theta_{[j]}-\theta+d_1) \; ds$$

$$+ \; m(m-1) \sum_{i=m+1}^{k} \int_{-\infty}^{\infty} \int_{-\infty}^{s+\theta\;-\theta_{[i]}+d_1} (F(s+d_1) - F(t+\theta_{[i]}-\theta-d_1))^{m-2}$$

$$\prod_{\substack{l=m+1 \\ l \neq i}}^{k} (F(s+\theta-\theta_{[l]}+d_1) - F(t+\theta_{[i]}-\theta_{[l]}-d_1))$$

$$(f(t+\theta_{[i]}-\theta-d_1)) \; f(s)f(t)dt\; ds$$

$$+ m \sum_{i=m+1}^{k} \int_{-\infty}^{\infty} \int_{-\infty}^{s+\theta-\theta_{[i]}+d_1} (F(s+d_1) - F(t+\theta_{[i]}-\theta-d_1))^{m-1}$$

$$\sum_{\substack{j=m+1 \\ j\neq i}}^{k} \prod_{\substack{l=m+1 \\ l\neq i,j}}^{k} (F(s+\theta-\theta_{[l]}+d_1) - F(t+\theta_{[i]}-\theta_{[l]}-d_1))$$

$$f(s+\theta-\theta_{[j]}+d_1)\ f(s)f(t)dt\ ds$$

$$+ m \sum_{i=m+1}^{k} \int_{-\infty}^{\infty} (F(s+d_1) - F(s)) \prod_{\substack{l=m+1 \\ l\neq i}}^{k} (F(s+\theta-\theta_{[l]}+d_1)$$

$$- F(s+\theta-\theta_{[l]}))\ f(s+\theta-\theta_{[i]}+d_1)\ f(s)ds$$

$$= \Big[m \sum_{i=m+1}^{k} \int_{-\infty}^{\infty} (F(s+d_1) - F(s))^{m-1} \prod_{\substack{l=m+1 \\ l\neq i}}^{k} (F(s+\theta-\theta_{[l]}+d_1)$$

$$- F(s+\theta-\theta_{[l]}))(f(s+\theta-\theta_{[i]}+d_1)f(s))-f(s+\theta-\theta_{[i]})\ f(s+d_1))\ ds\Big]$$

$$+ \Big[m \sum_{i=m+1}^{k} \sum_{\substack{j=m+1 \\ i\neq j}}^{k} \int_{-\infty}^{\infty} \int_{-\infty}^{s+d_1} (F(s+d_1) - F(t-d_1))^{m-1}$$

$$\prod_{\substack{l=m+1 \\ l\neq i,j}}^{k} (F(s+\theta-\theta_{[l]}+d_1) - F(t+\theta-\theta_{[l]}-d_1))$$

$$\times \{(f(t-d_1)-f(s+d_1))\ (f(s-\theta_{[j]}+\theta)f(t-\theta_{[i]}+\theta)) -$$

$$f(t+\theta-\theta_{[i]}-d_1)f(t)f(s-\theta_{[j]}+\theta)+f(s+\theta-\theta_{[j]}+d_1)f(s)f(t-\theta_{[i]}+\underset{\sim}{\theta})\}dt\ ds\Big]$$

$$+ \Bigg[m(m-1) \sum_{i=m+1}^{k} \int_{-\infty}^{\infty} \int_{-\infty}^{s+d_1} (F(s+d_1) - F(t-d_1))^{m-2}$$

$$\prod_{\substack{l=m+1 \\ l \neq i}}^{k} (F(s+\theta-\theta_{[l]}+d_1) - F(t+\theta-\theta_{[l]}-d_1))$$

$$(f(s+\theta-\theta_{[i]}+d_1) f(s) f(t) - f(t+\theta-\theta_{[i]}-d_1)\ f(s) f(t)$$

$$-f(s+d_1)\ f(s-\theta_{[i]}+\theta) f(t) + f(t-d_1) f(t-\theta_{[i]}+\theta)\ f(s))\ dt\ ds \Bigg]$$

Now on using MLR property we see that all the three terms are non-negative and thus,

$$\frac{d}{d\theta} Q_1(\theta) \geq 0 \qquad \forall\ \underset{\sim}{\theta} < \underset{\sim}{\theta}_{m+1}.$$

which proves the first assertion of the theorem. Second assertion follows by substituting $\theta_{[1]} = \ldots = \theta_{[k]}$ in (3.3.4). ■

As a consequence of above lemma, we have

<u>**Theorem 3.3.3:**</u> Suppose that loss function L is given by (3.3.1) and pdf's $f_{X_i}(x_i;\theta_i) = f(x_i-\theta_i)$, $i = 1,\ldots,k$ have MLR property, then the decision rule R_{MD} is minimax in $\mathcal{D}(P^*)$.

<u>**Proof:**</u> Clearly $R_{MD} \in \mathcal{D}(P^*)$.

Also,

$$\sup_{\Omega} R(\underset{\sim}{\theta}, R_{MD}) = k(k-1)P^*.$$

the minimax value. Hence R_{MD} is minimax in $\mathcal{D}(P^*)$. ■

Now suppose $X_1,\ldots,X_k$ are independent and X_i has a pdf $f_{X_i}(x_i;\theta_i) = \frac{1}{\theta_i} f\left(\frac{x_i}{\theta_i}\right)$, $x_i > 0$, $\theta_i > 0$. Consider the decision rule

R_M defined by :

R_M : Select π_i in S_L if and only if $x_i \le \frac{x_{[1]}}{d_2}$

and

Select π_j in S_U if and only if $x_j \ge d_2\, x_{[k]}$

where $d_2 \in (0,1)$ is chosen so that

$$\inf_{\Omega} P_{\underset{\sim}{\theta}}(CS|R_M) = P^* \qquad (3.3.5)$$

<u>**Lemma 3.3.2:**</u> Suppose $f_{X_i}(x_i;\theta_i) = \frac{1}{\theta_i} f\left(\frac{x}{\theta_i}\right)$, $x_i > 0$, $\theta_i > 0$, $i = 1,\dots,k$ and suppose pdf's $f_{X_i}(\cdot;\theta_i)$, $i = 1,\dots,k$ have MLR property. Then for the loss function given by (3.3.1),

(i) Infimum of $P_{\underset{\sim}{\theta}}(CS|R_M)$ is attained when $\theta_1 = \dots = \theta_k$,

(ii) Supremum of $R(\underset{\sim}{\theta},R_M)$ is attained when $\theta_1 = \dots = \theta_k$, and

(iii) $\sup_{\Omega} R(\underset{\sim}{\theta},R_M) = k(k-1)P^*$

<u>**Proof:**</u>

(i) $P_{\underset{\sim}{\theta}}(CS|R_M) = P_{\underset{\sim}{\theta}}\left(X_{(1)} \le \frac{X_{[1]}}{d_2},\ X_{(k)} \ge d_2\, X_{[k]}\right)$

$$= \int_0^{\infty} \int_0^{\frac{\theta_{[k]}}{\theta_{[1]}}\frac{s}{d_2}} \prod_{j=2}^{k-1}\left(F\left(\frac{s}{d_2}\frac{\theta_{[k]}}{\theta_{[j]}}\right) - F\left(td_2\frac{\theta_{[1]}}{\theta_{[j]}}\right)\right) f(s)f(t)\,dt\,ds$$

$$\ge \int_0^{\infty} \int_0^{\frac{s}{d_2}} \left(F\left(\frac{s}{d_2}\right) - F\left(td_2\right)\right)^{k-2} f(s)f(t)\,dt\,ds$$

and bound is attained when $\theta_{[1]} = \dots = \theta_{[k]}$. Hence

$P_{\underset{\sim}{\theta}}(CS|R_M)$ attains its infimum at $\theta_1 = \dots = \theta_k$ and

$$\inf_{\Omega} P_{\underset{\sim}{\theta}}(CS|R_M) = \int_0^{\infty} \int_0^{\frac{s}{d_2}} \left[F\left(\frac{s}{d_2}\right) - F\left(td_2\right)\right]^{k-2} f(s)f(t)\, dt\, ds$$

Proof of (ii) and (iii) is similar to that of Lemma 3.3.1. ∎

Following theorem is an immediate consequence of Lemma 3.3.2.

Theorem 3.3.4: Under the assumptions of Lemma 3.3.2, the decision rule R_M is minimax in $\mathcal{D}(P^*)$.

The following Theorem 3.3.5, provides a set of necessary conditions for minimaxity of a rule in $\mathcal{D}(P^*)$ with respect to the loss function given by (3.3.1).

Theorem 3.3.5: Suppose that loss function is given by (3.3.1) and let δ^* be a minimax decision rule in $\mathcal{D}(P^*)$. Suppose $P_{\underset{\sim}{\theta}}$(Select π_i in S_L and π_j in $S_U|\delta^*$) is continuous at $\underset{\sim}{\theta}_E \in \Omega_E$ for all $i \neq j$. Then for all $\underset{\sim}{\theta}_E \in \Omega_E$,

(i) $R(\underset{\sim}{\theta}_E, \delta^*) = k(k-1)P^* = \sup_{\Omega} R(\underset{\sim}{\theta}, \delta^*)$

(ii) $P_{\underset{\sim}{\theta}_E}$(Select π_i in S_L and π_j in $S_U|\delta^*) = P^* = \inf_{\Omega} P_{\underset{\sim}{\theta}}(CS|\delta^*)$,

$\forall\ i \neq j$

Proof: Fix $\underset{\sim}{\theta}_E \in \Omega_E$. As in proof of Theorem 3.3.1, it follows that

$$P_{\underset{\sim}{\theta}_E}(\text{Select } \pi_i \text{ in } S_L \text{ and } \pi_j \text{ in } S_U|\delta^*) \geq P^*,\ \forall\ i \neq j \qquad (3.3.6)$$

(i) Since the decision rule δ^* is minimax in $\mathcal{D}(P^*)$, it follows that

$$k(k-1)P^* = \sup_{\Omega} R(\underset{\sim}{\theta},\delta^*)$$
$$\geq R(\underset{\sim}{\theta}_E,\delta^*)$$
$$= \sum_{\substack{i=1}}^{k} \sum_{\substack{j=1 \\ i\neq j}}^{k} P_{\underset{\sim}{\theta}_E}(\text{Select } \pi_i \text{ in } S_L \text{ and } \pi_j \text{ in } S_U|\delta^*)$$
$$\geq k(k-1)P^* \qquad \text{(from (3.3.6))}$$

So all inequalities are equalities. Hence (i) is true.

(ii) From (3.3.6), we have

$$P_{\underset{\sim}{\theta}_E}(\text{Select } \pi_i \text{ in } S_L \text{ and } \pi_j \text{ in } S_U|\delta^*) \geq P^*, \quad \forall\, i \neq j.$$

Now suppose,

$$P_{\underset{\sim}{\theta}_E}(\text{Select } \pi_i \text{ in } S_L \text{ and } \pi_j \text{ in } S_U|\delta^*) > P^*, \text{ for some } i \neq j.$$

Then

$$\sup_{\Omega} R(\underset{\sim}{\theta},\delta^*) = R(\underset{\sim}{\theta}_E,\delta^*) \qquad \text{(from (i))}$$
$$= \sum_{\substack{i=1}}^{k} \sum_{\substack{j=1 \\ i\neq j}}^{k} P_{\underset{\sim}{\theta}_E}(\text{Select } \pi_i \text{ in } S_L \text{ and } \pi_j \text{ in } S_U|\delta^*)$$
$$> k(k-1)P^*$$

which is a contradiction. Hence, the result follows. ■

3.4 Minimax Rules Under Heteroscedasticity

Suppose that the observation X_i from π_i has a normal distribution with mean θ_i and known variance σ_i^2, $i = 1,\ldots,k$, ($k \geq 3$), and suppose that $X_1,\ldots,X_k$ are statistically independent. In

the case when our goal is to select two extreme populations associated with $\theta_{[1]}$ and $\theta_{[k]}$ our action apace is given by

$$\mathscr{A}^* = \{(i,j) : i \in \{1,\ldots,k\},\ j \in \{1,\ldots,k\}\}$$

where taking action (i,j) corresponds to the selection of π_i as LEP and the selection of π_j as UEP.

For convenience, we assume that there are no ties among θ_i's. We call the decision rule $\delta^{N,E}$, which selects the population associated with $X_{[1]}$ as the LEP and the population associated with $X_{[k]}$ as the UEP, as the natural decision rule. For given decision rule $\delta \in \mathscr{D}$, and $\underset{\sim}{x} \in \mathscr{X}$, let $\delta((i,j)|\underset{\sim}{x})$ denote the conditional probability of taking action (i,j) given $\underset{\sim}{X} = \underset{\sim}{x}$.

Suppose that the risk associated with a decision rule δ is given by :

$$R(\underset{\sim}{\theta},\delta) = 1-P_{\underset{\sim}{\theta}}(CS|\delta) \tag{3.4.1}$$

To find a minimax decision rule, we first find the minimax value of the problem. The following theorem gives the minimax value of the problem.

Theorem 3.4.1: For the problem at hand, the minimax value is $1-\frac{1}{k(k-1)}$, that is,

$$\inf_{\delta\in\mathscr{D}} \sup_{\Omega} R(\underset{\sim}{\theta},\delta) = 1 - \frac{1}{k(k-1)}.$$

Proof. Consider the no-data rule δ^o defined by

$$\delta^o(i,j|\underset{\sim}{x}) = \frac{1}{k(k-1)}\ , \qquad \forall\ i \neq j,\ \underset{\sim}{x} \in \mathscr{X}.$$

Then,

$$R(\underset{\sim}{\theta},\delta^o) = 1 - \frac{1}{k(k-1)}\ , \quad \forall\ \underset{\sim}{\theta} \in \Omega.$$

Since the no-data rule δ^0 has constant risk $1 - \frac{1}{k(k-1)}$, it suffices to find a sequence of prior distributions for $\underset{\sim}{\theta} = (\theta_1,\ldots,\theta_k)$, such that the sequence of associated Bayes risks tends to this value $1 - \frac{1}{k(k-1)}$ (See Berger (1985), p. 354). For this, let

$$d\tau_n(\theta_1,\ldots,\theta_k) = n^{\frac{k}{2}} \prod_{i=1}^{k} \phi(\sqrt{n}\,\theta_i)\, d\underset{\sim}{\theta}, \qquad n = 1,2,\ldots$$

be a sequence of prior distributions for $\underset{\sim}{\theta} = (\theta_1,\ldots,\theta_k)$. Then, given $\underset{\sim}{X} = \underset{\sim}{x}$, $\theta_1,\ldots,\theta_k$ are independently distributed with θ_i having a normal distribution with mean $\left(\frac{x_i}{n\sigma_i^2+1}\right)$ and variance $\frac{\sigma_i^2}{n\sigma_i^2+1}$, $i = 1,\ldots,k$. Thus, the Bayes rule δ^B, say, which minimizes the posterior expected loss, yields the posterior risk

$$R_n(\delta^B|\underset{\sim}{x}) = \min_{\substack{i,j=1,\ldots,k\\ i\neq j}} (1-P(\theta_i = \theta_{[1]} \text{ and } \theta_j = \theta_{[k]}|\underset{\sim}{X}=\underset{\sim}{x}))$$

$$= 1 - \max_{\substack{i,j=1,\ldots,k\\ i\neq j}} P(\theta_i = \theta_{[1]} \text{ and } \theta_j = \theta_{[k]}|\underset{\sim}{X}=\underset{\sim}{x})$$

Define,

$$Z_l = \frac{\theta_l - \frac{x_l}{(n\sigma_l^2+1)}}{\left(\frac{n\sigma_l^2}{n\sigma_l^2+1}\right)^{\frac{1}{2}}}, \quad l = 1,\ldots,k.$$

Then

$$R_n(\delta^B|\underset{\sim}{x}) = \max_{i\neq j}(1-P(\theta_i = \theta_{[1]} \text{ and } \theta_j = \theta_{[k]}|\underset{\sim}{X}=\underset{\sim}{x}))$$

$$= 1 - \max_{\substack{i,j=1,\dots,k \\ i\neq j}} P\Bigg[Z_i\left(\frac{\sigma_i^2}{n\sigma_i^2+1}\right)^{\frac{1}{2}} + \frac{x_i}{n\sigma_i^2+1} \le Z_l\left(\frac{\sigma_l^2}{n\sigma_l^2+1}\right)^{\frac{1}{2}} + \frac{x_l}{n\sigma_l^2+1} \le$$

$$Z_j\left(\frac{\sigma_j^2}{n\sigma_j^2+1}\right)^{\frac{1}{2}} + \frac{x_j}{n\sigma_j^2+1}, \; l = 1,\dots,k, \; l \neq i,j, \text{ and}$$

$$Z_i\left(\frac{\sigma_i^2}{n\sigma_i^2+1}\right)^{\frac{1}{2}} + \frac{x_i}{n\sigma_i^2+1} \le Z_j\left(\frac{\sigma_j^2}{n\sigma_j^2+1}\right)^{\frac{1}{2}} + \frac{x_j}{n\sigma_j^2+1}\Bigg]$$

Let

$$u_{n,i,j}(s) = \left(\frac{n\sigma_i^2+1}{\sigma_i^2}\right)^{\frac{1}{2}} \left[s\left(\frac{\sigma_j^2}{n\sigma_j^2+1}\right)^{\frac{1}{2}} + \frac{x_j}{n\sigma_j^2+1} - \frac{x_i}{n\sigma_i^2+1}\right]$$

so that

$R_n(\delta^B|\underset{\sim}{x})$

$$= 1 - \max_{\substack{i,j=1,\dots,k \\ i\neq j}} \int_{-\infty}^{\infty} \int_{-\infty}^{u_{n,i,j}(s)} \prod_{\substack{l=1 \\ l\neq i,j}}^{k} (\Phi(u_{n,l,j}(s)) - \Phi(u_{n,l,i}(t)) \times \phi(t)\phi(s)\, dt\, ds$$

Since

$$\lim_{n\to\infty} u_{n,i,j}(s) = s.$$

Therefore,

$$\lim_{n\to\infty} R_n(\delta^B|\underset{\sim}{x})$$

$$= 1 - \int_{-\infty}^{\infty} \int_{-\infty}^{s} (\Phi(s) - \Phi(t))^{k-2}\, \phi(s)\, \phi(t)\, dt\, ds$$

$$= 1 - \frac{1}{k(k-1)}.$$

For the prior τ_n $(n = 1,2,\ldots)$, let $r(\tau_n,\delta)$ denotes the Bayes risk associated with the decision rule δ. Then after repeated application of Lebesgue dominated convergence theorem, we get

$$\lim_{n\to\infty} r(\tau_n,\delta^B) = 1 - \frac{1}{k(k-1)}$$

This completes the proof of the theorem. ∎

The following lemma will be needed in examining the minimaxity of the natural decision rule $\delta^{N,E}$.

Lemma 3.4.1: Let $\gamma_1,\ldots,\gamma_{k-1}$ be k-1 positive constants. The the function

$$h(\gamma_1,\ldots,\gamma_{k-1}) = \int_{-\infty}^{\infty}\int_{-\infty}^{x\gamma_1} \prod_{j=2}^{k-1} \left[\Phi(x\gamma_j) - \Phi(y\frac{\gamma_j}{\gamma_1})\right] \phi(y)\,\phi(x)\,dy\,dx$$

is strictly decreasing in γ_1 and strictly increasing in γ_j, $j = 2,\ldots,k-1$.

Proof: Consider the partial derivative of h with respect to γ_2, keeping all other γ_i's fixed

$$\frac{\partial h}{\partial\gamma_2} = \int_{-\infty}^{\infty}\int_{-\infty}^{x\gamma_1} \prod_{j=3}^{k-1} \left[\Phi(x\gamma_j) - \Phi(y\frac{\gamma_j}{\gamma_1})\right]\left[x\,\phi(x\gamma_2) - \frac{y}{\gamma_1}\,\phi(y\,\frac{\gamma_2}{\gamma_1})\right] \phi(y)\,\phi(x)\,dy\,dx$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{x\gamma_1} x \prod_{j=3}^{k-1} \left[\Phi(x\gamma_j) - \Phi(y\frac{\gamma_j}{\gamma_1})\right] \phi(y)\,\phi(x)\,\phi(x\gamma_2)\,dy\,dx$$

$$- \int_{-\infty}^{\infty}\int_{-\infty}^{x\gamma_1} \frac{y}{\gamma_1} \prod_{j=3}^{k-1} \left[\Phi(x\gamma_j) - \Phi(y\frac{\gamma_j}{\gamma_1})\right] \phi(y\,\frac{\gamma_2}{\gamma_1})\,\phi(y)\,\phi(x)\,dy\,dx \tag{3.4.3}$$

$$= C_1 - D_1, \text{ say.}$$

where,

$$C_1 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{x\gamma_1} x \prod_{j=3}^{k-1} \left[\Phi(x\gamma_j) - \Phi(y \frac{\gamma_j}{\gamma_1})\right] \phi(x \sqrt{1+\gamma_2^2})\, \phi(y)\, dy\, dx$$

and

$$D_1 = \frac{1}{\gamma_1} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{x\gamma_1} y\phi\left(y\sqrt{1 + \frac{\gamma_2^2}{\gamma_1^2}}\right) \prod_{j=3}^{k-1} \left[\Phi(x\gamma_j) - \Phi(y \frac{\gamma_j}{\gamma_1})\right]\Big]$$

$$\phi(y)\, \phi(x)\, dy\, dx$$

Now

$$C_1 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{x\gamma_1} x \prod_{j=3}^{k-1} \left[\Phi(x\gamma_j) - \Phi(y \frac{\gamma_j}{\gamma_1})\right] \phi(x\sqrt{1+\gamma_2^2})\, \phi(y)\, dy \quad dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left\{ \int_{\frac{y}{\gamma_1}}^{\infty} x\, \phi(x \sqrt{1+\gamma_2^2}) \prod_{j=3}^{k-1} \left[\Phi(x\gamma_j) - \Phi(y \frac{\gamma_j}{\gamma_1})\right] \phi(y) \quad dy\ dx\right.$$

On integrating the inside integral by parts, we get

$$C_1 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[- \frac{\phi(x\sqrt{1+\gamma_2^2})}{1+\gamma_2^2} \prod_{j=3}^{k-1} \left[\Phi(x\gamma_j) - \Phi(y \frac{\gamma_j}{\gamma_1})\right]\right]_{\frac{y}{\gamma_1}}^{\infty} \phi(y)\, dy$$

$$+ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{\frac{y}{\gamma_1}}^{\infty} \frac{\partial}{\partial x} \left[\prod_{j=3}^{k-1} \left[\Phi(x\gamma_j) - \Phi(y \frac{\gamma_j}{\gamma_1})\right]\right] \phi\left(\frac{x\sqrt{1+\gamma_2^2}}{1+\gamma_2^2}\right) \phi(y)\, dx\, dy$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{\frac{y}{\gamma_1}}^{\infty} \sum_{i=3}^{k-1} \prod_{\substack{j=3 \\ j\neq i}}^{k-1} \left[\Phi(x\gamma_j) - \Phi\left(y \frac{\gamma_j}{\gamma_1}\right)\right] \gamma_i \, \phi(x\gamma_i)$$

$$\phi\left[\frac{x\sqrt{1+\gamma_2^2}}{1+\gamma_2^2}\right] \phi(x) \, dx \, dy$$

$$> 0. \tag{3.4.4}$$

$$D_1 = \frac{1}{\sqrt{2\pi}} \frac{1}{\gamma_1} \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{x\gamma_1} y \, \phi\left(y \sqrt{1+\gamma_2^2}\right) \prod_{j=3}^{k-1} \left[\Phi(x\gamma_j) - \Phi\left(y \frac{\gamma_j}{\gamma_1}\right)\right] dy \right\} \phi(x) dx$$

On integrating the inside integral by parts, we get

$$D_1 = \frac{1}{\gamma_1} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[- \frac{\phi\left[y\sqrt{1 + \frac{\gamma_2^2}{\gamma_1^2}}\right]}{1 + \frac{\gamma_2^2}{\gamma_1^2}} \prod_{j=3}^{k-1} \left[\Phi(x\gamma_j) - \Phi\left(y \frac{\gamma_j}{\gamma_1}\right)\right]\right]_{-\infty}^{x\gamma_1} \phi(x) \, dx$$

$$+ \frac{1}{\gamma_1} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{x\gamma_1} \frac{\phi\left[y\sqrt{1 + \frac{\gamma_2^2}{\gamma_1^2}}\right]}{1 + \frac{\gamma_2^2}{\gamma_1^2}} \frac{\partial}{\partial y} \prod_{j=3}^{k-1} \left[\Phi(x\gamma_j) - \Phi\left(y \frac{\gamma_j}{\gamma_1}\right)\right] \phi(x) \, dy dx$$

$$= \frac{-\gamma_1}{\gamma_1^2 + \gamma_2^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{x\gamma_1} \left[\sum_{i=3}^{k-1} \prod_{\substack{j=3 \\ j\neq i}}^{k-1} \left[\Phi(x\gamma_j) - \Phi\left(y \frac{\gamma_j}{\gamma_1}\right)\right] \frac{\gamma_i}{\gamma_1} \phi\left(y \frac{\gamma_i}{\gamma_1}\right)\right]$$

$$\phi\left[y \sqrt{1 + \frac{\gamma_2^2}{\gamma_1^2}}\right] \phi(x) \, dy \, dx.$$

$$< 0 \tag{3.4.5}$$

On combining (3.4.3) - (3.4.5), we get

$$\frac{\partial h}{\partial \gamma_2} > 0$$

Since h is symmetric function of $\gamma_2, \ldots, \gamma_{k-1}$,

$$\frac{\partial h}{\partial \gamma_j} > 0, \text{ for } j = 2, \ldots, k-1$$

and hence, h is strictly increasing function of $\gamma_2, \ldots, \gamma_{k-1}$.

Now consider the partial derivative of h with respect to γ_1,

$$\frac{\partial h}{\partial \gamma_1} = \int_{-\infty}^{\infty} x \prod_{j=2}^{k-1} (\Phi(x\gamma_j) - \Phi(x\gamma_j))\, \phi(x\gamma_1)\, \phi(x) dx$$

$$+ \int_{-\infty}^{\infty} \int_{-\infty}^{x\gamma_1} \sum_{i=2}^{k-1} \prod_{\substack{j=2 \\ j\neq i}}^{k-1} \left[\Phi(x\gamma_j) - \Phi(y \frac{\gamma_j}{\gamma_1})\right] y \frac{\gamma_i}{\gamma_1^2} \phi(y \frac{\gamma_i}{\gamma_1})\, \phi(y)\phi(x)\ dy\ dx$$

$$= \frac{1}{\gamma_1^2} \frac{1}{\sqrt{2\pi}} \sum_{i=2}^{k-1} \gamma_i \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{x\gamma_1} y\ \phi(y \sqrt{1+\frac{\gamma_i^2}{\gamma_1^2}}) \prod_{\substack{j=2 \\ j\neq i}}^{k-1} \left[\Phi(x\gamma_j) - \Phi(y \frac{\gamma_j}{\gamma_1})\right] dy \right\} \phi(x) dx$$

On integrating the inside integral by parts we get

$$\frac{\partial h}{\partial \gamma_1} = \frac{1}{\gamma_1^2} \frac{1}{\sqrt{2\pi}} \sum_{i=2}^{k-1} \gamma_i \int_{-\infty}^{\infty} \left[\frac{\phi\left(y \sqrt{1 + \frac{\gamma_i^2}{\gamma_1^2}}\right)}{1 + \frac{\gamma_i^2}{\gamma_1^2}} \prod_{\substack{j=2 \\ j\neq i}}^{k-1} \left[\Phi(x\gamma_j) - \Phi(y \frac{\gamma_j}{\gamma_1})\right]\right]_{-\infty}^{x\gamma_1} \phi(x) dx$$

$$+ \frac{1}{\gamma_1^2} \frac{1}{\sqrt{2\pi}} \sum_{i=2}^{k-1} \gamma_i \int_{-\infty}^{\infty} \int_{-\infty}^{x\gamma_1} \frac{\partial}{\partial y} \left(\prod_{\substack{j=2 \\ j\neq i}}^{k-1} \left[\Phi(x\gamma_j) - \Phi(y \frac{\gamma_j}{\gamma_1})\right]\right) \frac{\phi\left[y \sqrt{1 + \frac{\gamma_i^2}{\gamma_1^2}}\right]}{1 + \frac{\gamma_i^2}{\gamma_1^2}} \phi(y)\phi(x) dy dx$$

$$= \frac{1}{\sqrt{2\pi}} \sum_{i=2}^{k-1} \frac{\gamma_i}{\gamma_1^2 + \gamma_i^2} \int_{-\infty}^{\infty} \int_{-\infty}^{x\gamma_1} \frac{\partial}{\partial y} \left(\prod_{\substack{j=2 \\ j \neq i}}^{k-1} \left[\Phi(x\gamma_j) - \Phi(y \frac{\gamma_j}{\gamma_1}) \right] \right)$$

$$\phi(y \sqrt{1 + \frac{\gamma_i^2}{\gamma_1^2}}) \, \phi(x) \, dy \, dx$$

$$< 0$$

and the result follows. ∎

Now we prove the main result of this section.

Theorem 3.4.2: The natural decision rule $\delta^{N,E}$ is minimax if and only if $\sigma_1^2 = \ldots = \sigma_k^2$.

Proof: Let $\sigma_{(i)}^2$ denotes the variance of the population associated with $\theta_{[i]}$ and let

$$v_{i,j}(x) = \frac{x\sigma_{(j)} + \theta_{[j]} - \theta_{[i]}}{\sigma(i)}, \qquad i,j = 1,\ldots,k.$$

Then

$$R(\underset{\sim}{\theta}, \delta^{N,E}) = 1 - P_{\underset{\sim}{\theta}}(X_{(1)} = X_{[1]} \text{ and } X_{(k)} = X_{[k]})$$

$$= 1 - \int_{-\infty}^{\infty} \int_{-\infty}^{v_{1,k}(x)} \prod_{j=2}^{k} (\Phi(v_{j,k}(x)) - \Phi(v_{j,1}(y)) \, \phi(y) \, \phi(x) dy \, dx$$

Thus,

$$\sup_{\Omega} R(\underset{\sim}{\theta}, \delta^{N,E})$$

$$= 1 - \inf_{\Omega} \int_{-\infty}^{\infty} \int_{-\infty}^{v_{1,k}(x)} \prod_{j=2}^{k-1} (\Phi(v_{j,k}(x) - \Phi(v_{j,1}(x))) \, \phi(y) \, \phi(x) dy \, dx$$

$$= 1 - \int_{-\infty}^{\infty} \int_{-\infty}^{\frac{x\sigma_{(k)}}{\sigma_{(1)}}} \prod_{j=2}^{k-1} \left(\Phi\left(x \frac{\sigma_{(k)}}{\sigma_{(j)}}\right) - \Phi\left(y \frac{\sigma_{(1)}}{\sigma_{(j)}}\right)\right) \phi(y)\phi(x)\, dy\, dx$$

$$\geq 1 - \int_{-\infty}^{\infty} \int_{-\infty}^{\frac{x\sigma_{[2]}}{\sigma_{[1]}}} \prod_{j=2}^{k-1} \left(\Phi\left(x \frac{\sigma_{[2]}}{\sigma_{[j+1]}}\right) - \Phi\left(y \frac{\sigma_{[1]}}{\sigma_{[j+1]}}\right)\right) \phi(y)\phi(x)\, dy\, dx$$

$$= 1 - \int_{-\infty}^{\infty} \int_{-\infty}^{\frac{x\sigma_{[2]}}{\sigma_{[1]}}} \prod_{j=2}^{k-1} \left(\Phi\left(x \frac{\sigma_{[2]}}{\sigma_{[j+1]}}\right) - \Phi\left(y \frac{\sigma_{[2]}/\sigma_{[j+1]}}{\sigma_{[2]}/\sigma_{[1]}}\right)\right) \phi(y) d(x)\, dy\, dx \qquad (3.4.6)$$

$$\geq 1 - \int_{-\infty}^{\infty} \int_{-\infty}^{x} (\Phi(x) - \Phi(y))^{k-2} \phi(y)\, \phi(x)\, dy\, dx$$

(from Lemma 3.4.1)

$$= 1 - \frac{1}{k(k-1)} .$$

With equality holding if and only if $\sigma_1^2 = \dots = \sigma_k^2$ (in view of Lemma 3.4.1). Thus, the result follows from Theorem 3.4.1. ■

We conclude this section with the following remarks.

__Remark 1.__ From (3.4.6) we can see if the parameters $\theta_1, \dots, \theta_k$ are close together, and if variances $\sigma_{(k)}^2$ and $\sigma_{(1)}^2$ are relatively small in comparison with other variances and simultaneously if $\sigma_{(k)}^2$ is relatively larger than $\sigma_{(1)}^2$, then the rule $\delta^{N,E}$ performs worse than the rule which selects at random. Dudewicz and Dalal (1975) also obtained a similar counterintuitive result.

__Remark 2__: Results of this section extend to the case of independent samples of sizes $n_1, \dots, n_k$ from populations $\pi_1, \dots, \pi_k$ with known variances $\sigma_1^2, \dots, \sigma_k^2$ by taking $x_1, \dots, x_k$ as sample means and σ_i^2/n_i, $i=1,\dots,k$, as the known variances.

CHAPTER IV

SIMULTANEOUS SELECTION OF EXTREME POPULATIONS: OPTIMAL TWO-STAGE PROCEDURES

For the problem of selecting the UEP in two stages, Gupta and Miescke (1984) have derived an essentially complete class of two-stage procedures for a general loss function. In this chapter we consider the goal of simultaneously selecting the LEP and the UEP in two stages: screening out non-extreme populations at the first stage and selecting two populations associated with $\theta_{[1]}$ and $\theta_{[k]}$ at the second stage. In Section 4.1 we formulate the problem and give the definition of two-stage permutation invariant procedures. In Sections 4.2 and 4.3 we derive two-stage Bayes and two-stage permutation invariant procedures.

4.1 Formulation of Problem

Suppose that the observation X_i from π_i has a normal distribution with mean θ_i $(-\infty < \theta_i < \infty)$ and common known variance σ^2(=1, without loss of generality), $i = 1,\ldots,k$, $k \geq 3$, and suppose that r.v.'s $X_1,\ldots,X_k$ are independent. We are interested in deriving an optimal two-stage procedure for the problem of simultaneously selecting the LEP and the UEP. We will consider the following class of two-stage procedures :

At stage 1, observe $X_{i1},\ldots,X_{in_1}$ from π_i, $i = 1,\ldots,k$, and based on these independent observations select two non-empty and

disjoint subsets S_L and S_U. If $|S_L| = |S_U| = 1$, then procedure stops at stage 1 and claims that populations is S_L and S_U correspond to $\theta_{[1]}$ and $\theta_{[k]}$, respectively. In all other cases procedure proceeds to stage 2.

If at stage 1, $|S_L| = 1$, $2 \le |S_U| \le k-1$ ($|S_U| = 1$, $2 \le |S_L| \le k-1$,), $S_L \cap S_U = \phi$, then procedure claims that population in $S_L(S_U)$ correspond to $\theta_{[1]}$ $(\theta_{[k]})$ and decision about population corresponding to $\theta_{[k]}$ $(\theta_{[1]})$ is made after taking n_2 additional observations $Y_{i1},\ldots,Y_{in_2}$ from each π_i, $i \in S_U$ ($i \in S_L$).

If at stage 1, $2 \le |S_L|$, $|S_U| \le k$, $4 \le |S_L| + |S_U| \le k$, $S_L \cap S_U = \phi$, then n_2 additional independent observations are taken from each π_i, $i \in S_L \cup S_U$ and final decision is based on two samples from each π_i, $i \in S_L \cup S_U$. Let $U_i = X_{i1} + \ldots + X_{in_1}$ and $V_i = Y_{i1} + \ldots + Y_{in_2}$, $i = 1,\ldots,k$ be sufficient statistics for θ_i at stage 1 and stage 2, respectively, and suppose that $f(.;\theta_i)$ and $h(.;\theta_i)$ denote the pdf's of U_i and V_i, $i = 1,\ldots,k$, respectively. Thus

$$f(u_i;\theta_i) = \frac{1}{\sqrt{2\pi n_1}} \exp\left[-\frac{1}{2n_1}(u_i - n_1\theta_i)^2 \right], \qquad i = 1,\ldots,k$$

and

$$h(u_i;\theta_i) = \frac{1}{\sqrt{2\pi n_2}} \exp\left[-\frac{1}{2n_2}(u_i - n_1\theta_i)^2 \right], \qquad i = 1,\ldots,k$$

Let $W_i = U_i + V_i$, $\underset{\sim}{U} = (U_1,\ldots,U_k)$, $\underset{\sim}{V} = (V_1,\ldots,V_k)$ and $\underset{\sim}{W} = (W_1,\ldots,W_k)$ be the joint sufficient statistics for $\underset{\sim}{\theta} = (\theta_1,\ldots,\theta_k)$. We will consider only those decision rules which

depend on observations only through sufficient statistics. In order to define a two-stage decision rule, let

$$\nu = \{\nu_{s,t} \mid s,t = 1,\dots,k,\ 2 \le s+t \le k\}$$

where for each $s,t = 1,\dots,k$, $2 \le s+t \le k$, $\nu_{s,t}$ is a measurable map from $\mathbb{R}^k$ to $[0,1]$ with the interpretation: after observing $\underset{\sim}{U} = \underset{\sim}{u}$, $\nu_{s,t}(\underset{\sim}{u})$ denotes the conditional probability of selecting s populations in S_L and t populations in S_U at stage 1, given $\underset{\sim}{U} = \underset{\sim}{u}$. Let

$$\eta(\underset{\sim}{u}) = \{\eta_{s,t}((a_1,a_2)\mid \underset{\sim}{u}) : a_1,a_2 \subset \{1,\dots,k\},\ a_1 \cap a_2 = \phi,$$
$$|a_1| = s,\ |a_2| = t,\ s,t = 1,\dots,k,\ 2 \le s+t \le k\},$$

where $\eta_{s,t}$ is a measurable map from $\mathbb{R}^k$ to $[0,1]$ with the interpretation that $\eta_{s,t}((a_1,a_2)\mid \underset{\sim}{u})$ denotes the conditional probability of selecting subset pair (a_1,a_2) with $|a_1| = s$, $|a_2| = t$, having observed $\underset{\sim}{U} = \underset{\sim}{u}$ at stage 1.

If $\nu(\underset{\sim}{u})$ decides that $|S_L| = |S_U| = 1$, then procedure stops at stage 1, and a final decision is based on

$$\{\eta_{1,1}((\{i\},\{j\})\mid \underset{\sim}{u}),\ i,j = 1,\dots,k,\ i \ne j\}$$

In all other cases procedure proceeds to stage 2. If at stage 1 η selects $S_L = \{i\}$, $S_U = \{j_1,\dots,j_t\}$ with $j_1 < \dots < j_t$, $2 \le t \le k-1$, $S_L \cap S_U = \phi$ ($S_L = \{i_1,\dots,i_s\}$, $S_U = \{j\}$ with $i_1<\dots<i_s$, $2 \le s \le k-1$, $S_L \cap S_U = \phi$) then after having observed $V_{j_1} = v_{j_1},\dots,V_{j_t} = v_{j_t}$ ($V_{i_1} = v_{i_1},\dots,V_{i_s} = v_{i_s}$) at stage 2, final decision is based on $\{\delta_{i,j,\{i\},S_U}(\underset{\sim}{u};v_{j_1},\dots,v_{j_t}) \mid j \in S_U\}$ ($\{\delta_{i,j,S_L,\{j\}}(\underset{\sim}{u};v_{i_1},\dots,v_{i_s}) \mid i \in S_L\}$). On the other hand if at stage 1 η selects $S_L = \{i_1,\dots,i_s\}$ and $S_U = \{j_1,\dots,j_t\}$, $S_L \cap S_U = \phi$ with $i_1 < \dots < i_s$,

$j_1 < \dots < j_t$, $s,t \geq 2$, $4 \leq s+t \leq k$, then after having observed $V_{i_1} = v_{i_1}, \dots, V_{i_s} = v_{i_s}$; $V_{j_1} = v_{j_1}, \dots, V_{j_t} = v_{j_t}$, the final decision is based on

$$\{\delta_{i,j,S_L,S_U}(\underset{\sim}{u}; v_{i_1}, \dots, v_{i_s}; v_{j_1}, \dots, v_{j_t}) \mid i \in S_L,\ j \in S_U\}$$

where

$$\delta = \{\delta_{i,j,a_1,a_2} \mid i \in a_1,\ j \in a_2,\ a_1, a_2 \subset \{1,\dots,k\},$$
$$a_1 \cap a_2 = \phi,\ 3 \leq |a_1| + |a_2| \leq k\}$$

δ_{i,j,a_1,a_2} is a measurable map from $\mathbb{R}^k \times \mathbb{R}^{|a_1|+|a_2|}$ to $[0,1]$ such that

$$\sum_{i \in a_1} \sum_{j \in a_2} \delta_{i,j,a_1,a_2} \equiv 1.$$

Definition 4.1.1 : The triple (ν,η,δ) will be referred to as two-stage decision rule (procedure), where ν, η, δ are as explained above.

Let $\mathfrak{D}$ denote the class of all two-stage procedures.

Definition 4.1.2 : We will say that a two-stage procedure (ν,η,δ) is permutation invariant if

(i) $\nu_{s,t}(\underset{\sim}{u}) = \nu_{s,t}(g\underset{\sim}{u})$, $\forall\ g \in G$, $\forall\ s,t \in \{1,\dots,k\}$, $2 \leq s+t \leq k$,

(ii) $\forall\ s,t \in \{1,\dots,k\}$, $2 \leq s+t \leq k$, $a_1, a_2 \subset \{1,\dots,k\}$, $a_1 \cap a_2 = \phi$ with $|a_1| = s$, $|a_2| = t$,

$$\eta_{s,t}((a_1,a_2) \mid \underset{\sim}{u}) = \eta_{s,t}((ga_1, ga_2) \mid g\underset{\sim}{u}),\ \forall\ g \in G.$$

(iii) for $a_1 = \{i_1,\dots,i_s\}$, $a_2 = \{j_1,\dots,j_t\}$ with $i_1 < \dots < i_s$ and $j_1 < \dots < j_t$; $ga_1 = \{i'_1,\dots,i'_s\}$, $ga_2 = \{j'_1,\dots,j'_t\}$ with $i'_1 < \dots < i'_s$ and $j'_1 < \dots < j'_t$,

$$\delta_{i,j,a_1,a_2}(\underset{\sim}{u}; v_{gi_1}, \dots, v_{gi_s}; v_{gj_1}, \dots, v_{gj_t})$$

$$= \delta_{gi,gj,ga_1,ga_2}(g\underset{\sim}{u}; v'_{i_1},\ldots,v'_{i_s}; v'_{j_1},\ldots,v'_{j_t})$$

Let $\mathfrak{D}_I$ denote the class of two-stage permutation invariant procedures. Let τ be a symmetric prior on $\Omega = \mathbb{R}^k$. Let $L(\underset{\sim}{\theta},(a_1,a_2),i,j)$ denote the loss incurred in selecting subset pair (a_1,a_2) at stage 1 and then finally selecting populations π_i and π_j as the LEP and the UEP, respectively. Assume that

$$L(\underset{\sim}{\theta},(a_1,a_2),i,j) = L(g\underset{\sim}{\theta},(ga_1,ga_2),\ gi,gj),\ \forall\ g \in G \tag{4.1.1}$$

In addition to (4.1.1), suppose that the loss function satisfies

$$\left.\begin{array}{ll}
\text{(a)} & L(\underset{\sim}{\theta},(\{i\},\{j\}),i,j) \le L(\underset{\sim}{\theta},(\{i'\},\{j\}),i',j),\ \text{for} \\
& \qquad \theta_i \le \theta_{i'},\ i \ne j,\ i' \ne j. \\
\text{(b)} & L(\underset{\sim}{\theta},(\{i\},\{j\}),i,j) \le L(\underset{\sim}{\theta},(\{i\},\{j'\}),i,j'),\ \text{for} \\
& \qquad \theta_j \ge \theta_{j'},\ i \ne j,\ i \ne j' \\
\text{(c)} & L(\underset{\sim}{\theta},(a_1,a_2),i,j) \le L(\underset{\sim}{\theta},(a_1,a_2),i',j),\ \text{for} \\
& \qquad \theta_i \le \theta_{i'},\ i,i' \in a_1,\ j \in a_2,\ a_1 \cap a_2 = \phi \\
\text{(d)} & L(\underset{\sim}{\theta},(a_1,a_2),i,j) \le L(\underset{\sim}{\theta},(a_1,a_2),i,j'),\ \text{for} \\
& \qquad \theta_j \ge \theta_{j'},\ i \in a_1,\ j,\ j' \subset a_2,\ a_1 \cap a_2 = \phi \\
\text{(e)} & \text{For } a_1 = \{i_1,\ldots,i_{r-1},i\}, a_2 = \{j_1,\ldots,j_{s-1},j\}, \tilde{a}_1 = \\
& \{i_1,\ldots,i_{r-1},i'\}, \tilde{a}_2 = \{j_1,\ldots,j_{s-1},j'\} \text{ with } a_1 \cap a_2 = \phi, \\
& \tilde{a}_1 \cap \tilde{a}_2 = \phi,\ a_1 \cap \tilde{a}_2 = \phi,\ \tilde{a}_1 \cap a_2 = \phi, \\
\text{(i)} & L(\underset{\sim}{\theta},(a_1,a_2),q,m) \le L(\underset{\sim}{\theta},(\tilde{a}_1,a_2),q,m), \text{for } \theta_{i'} > \theta_i \\
\text{(ii)} & L(\underset{\sim}{\theta},(a_1,a_2),q,m) \le L(\underset{\sim}{\theta},(a_1,\tilde{a}_2),q,m), \text{for } \theta_i \ge \theta_{i'} \\
\text{(iii)} & L(\underset{\sim}{\theta},(a_1,a_2),i,m) \le L(\underset{\sim}{\theta},(\tilde{a}_1,a_2),i',m), \text{for } \theta_i \le \theta_{i'} \\
\text{(iv)} & L(\underset{\sim}{\theta},(a_1,a_2),q,j) \le L(\underset{\sim}{\theta},(a_1,\tilde{a}_2),q,j'), \text{for } \theta_j \ge \theta_{j'}
\end{array}\right\} \tag{4.1.2}$$

Under the above set up we will first find optimal second stage decision rule δ and then the optimal two-stage decision rule.

4.2 The Optimal Second Stage Decision Rule

For every $s,t \in \{1,\ldots,k\}$, $a_1,a_2 \subset \{1,\ldots,k\}$, $a_1 \cap a_2 = \phi$ with $|a_1| = s$, $|a_2| = t$ and $\underset{\sim}{u} \in \mathbb{R}^k$, define

$$\eta^*_{s,t}((a_1,a_2)|\underset{\sim}{u}) = \begin{cases} \dfrac{1}{|H_{\underset{\sim}{u}}(s,t)|}, & \text{if } (a_1,a_2) \in H_{\underset{\sim}{u}}(s,t) \\ 0, & \text{otherwise} \end{cases}$$

where

$$H_{\underset{\sim}{u}}(s,t) = \{(a_1,a_2) : |a_1| = s, |a_2| = t,\ a_1 \cap a_2 = \phi,\ u_i \le u_q \le u_j,\ \forall\ i \in a_1,\ j \in a_2,\ q \in a_1^c \cap a_2^c\}.$$

Similarly for $a_1 = \{i_1,\ldots,i_s\}$, $a_2 = \{j_1,\ldots,j_t\}$ with $i_1 <\ldots< i_s$; $j_1 <\ldots< j_t$, $a_1 \cap a_2 = \phi$, $s,t \ge 1$, $\underset{\sim}{u} \in \mathbb{R}^k$, $(v_{i_1},\ldots,v_{i_s}) \in \mathbb{R}^s$, $(v_{j_1},\ldots,v_{j_t}) \in \mathbb{R}^t$, $w_l = u_l + v_l$, $l \in a_1 \cup a_2$, define

$$\delta^*_{i,j,a_1,a_2}(\underset{\sim}{u};v_{i_1},\ldots,v_{i_s};\ v_{j_1},\ldots,v_{j_t})$$

$$= \begin{cases} \dfrac{1}{|C_{s,t,a_1,a_2}(w_{i_1},\ldots,w_{i_s};w_{j_1},\ldots,w_{j_t})|}, \\ \qquad \text{if } (i,j) \in C_{s,t,a_1,a_2}(w_{i_1},\ldots,w_{i_s};w_{j_1},\ldots,w_{j_t}) \\ 0, \qquad \text{otherwise} \end{cases} \tag{4.2.1}$$

where

$$C_{s,t,a_1,a_2}(w_{i_1},\ldots,w_{i_s};w_{j_1},\ldots,w_{j_t}) = \{(i,j):w_i = \min\{w_q:q \in a_1\} \text{ and } w_j = \max\{w_q:q \in a_2\}\}$$

Now, the risk of a procedure (ν,η,δ) at $\underset{\sim}{\theta} \in \Omega$ is given by

$R(\underset{\sim}{\theta},(\nu,\eta,\delta))$

$$= \sum_{\substack{i=1\\ i\neq j}}^{k} \sum_{j=1}^{k} L(\underset{\sim}{\theta},(\{i\},\{j\}),i,j)\, E_{\underset{\sim}{\theta}} [\nu_{1,1}(\underset{\sim}{U})\eta_{1,1}((\{i\},\{j\})|\underset{\sim}{U})]$$

$$+ \sum_{i=1}^{k} \sum_{t=2}^{k-1} E_{\underset{\sim}{\theta}} [\nu_{1,t}(\underset{\sim}{U}) \sum_{\substack{a_2:a_2=\{j_1,\ldots,j_t\}\\ j_q\neq i,q=1,\ldots,t}} \eta_{1,t}((\{i\},a_2)|\underset{\sim}{U})$$

$$\sum_{j\in a_2} L(\underset{\sim}{\theta},(\{i\},a_2),\ i,j)\ \delta_{i,j,\{i\},a_2}(\underset{\sim}{U};V_{j_1},\ldots,V_{j_t})]$$

$$+ \sum_{s=1}^{k-1} \sum_{j=1}^{k} E_{\underset{\sim}{\theta}} [\nu_{s,1}(\underset{\sim}{U}) \sum_{\substack{a_1:a_1=\{i_1,\ldots,i_s\}\\ i_q\neq j,q=1,\ldots,s}} \eta_{s,1}((a_1\{j\})|\underset{\sim}{U})$$

$$\sum_{i\in a_1} L(\underset{\sim}{\theta},(a_1,\{j\}),i,j)\ \delta_{i,j,a_1,\{j\}}(\underset{\sim}{U};V_{i_1},\ldots,V_{i_s})]$$

$$+ \sum_{\substack{s=2\\ s+t\leq k}}^{k} \sum_{t=2}^{k} E_{\underset{\sim}{\theta}} [\nu_{s,t}(\underset{\sim}{U}) \sum_{\substack{(a_1,a_2):a_1=\{i_1,\ldots,i_s\}\\ a_2=\{j_1,\ldots,j_t\},a_1\cap a_2=\phi}} \eta_{s,t}((a_1,a_2)|\underset{\sim}{U})$$

$$\sum_{i\in a_1} \sum_{j\in a_2} L(\underset{\sim}{\theta},(a_1,a_2),i,j)$$

$$\delta_{i,j,a_1,a_2}(\underset{\sim}{U};V_{i_1},\ldots,V_{i_s};V_{j_1},\ldots,V_{j_t})]$$

The Bayes risk for (ν,η,δ) under prior τ is given by

$$r(\tau,(\nu,\eta,\delta)) = \int_{\Omega} R(\underset{\sim}{\theta},(\nu,\eta,\delta))\ d\tau(\underset{\sim}{\theta}) \tag{4.2.2}$$

Let $r^*_{\underset{\sim}{u}}((\nu,\eta,\delta))$ denotes the conditional posterior risk of a procedure (ν,η,δ) given $\underset{\sim}{U} = \underset{\sim}{u}$. In order to find a rule which minimizes (4.2.2), it is sufficient to find a rule which minimizes $r^*_{\underset{\sim}{u}}((\nu,\eta,\delta))$ for every $\underset{\sim}{u} \in \mathbb{R}^k$. Now,

$$r^*_{\underset{\sim}{u}}((\nu,\eta,\delta))$$

$$= D_1(\underset{\sim}{u},(\nu,\eta,\delta))+D_2(\underset{\sim}{u},(\nu,\eta,\delta))+D_3(\underset{\sim}{u},(\nu,\eta,\delta))+D_4(\underset{\sim}{u},(\nu,\eta,\delta)) \tag{4.2.3}$$

where

$$D_1(\underset{\sim}{u},(\nu,\eta,\delta))$$

$$= \nu_{1,1}(\underset{\sim}{u}) \sum_{\substack{i=1 \\ i\neq j}}^{k} \sum_{j=1}^{k} \nu_{1,1}((\{i\},\{j\})|\underset{\sim}{u}) E\big[L(\underset{\sim}{\theta},(\{i\},\{j\}),i,j)|\underset{\sim}{U}=\underset{\sim}{u}\big],$$

$$D_2(\underset{\sim}{u},(\nu,\eta,\delta))$$

$$= \sum_{i=1}^{k} \sum_{t=1}^{k-1} \nu_{1,t}(\underset{\sim}{u}) \sum_{\substack{a_2:a_2=\{j_1,\ldots,j_t\} \\ j_q\neq i,\, q=1,\ldots,t}} \eta_{1,t}((\{i\},a_2)|\underset{\sim}{u})$$

$$E\Big[\sum_{j\in a_2} \delta_{i,j,\{i\},a_2}(\underset{\sim}{u};V_{j_1},\ldots,V_{j_t})\, L(\underset{\sim}{\theta},(\{i\},a_2),i,j)|\underset{\sim}{U}=\underset{\sim}{u}\Big],$$

$$D_3(u,(\nu,\eta,\delta))$$

$$= \sum_{s=2}^{k} \sum_{j=1}^{k-1} \nu_{s,1}(\underset{\sim}{u}) \sum_{\substack{a_1:a_1=\{i_1,\ldots,i_s\} \\ i_q\neq j,\, q=1,\ldots,s}} \eta_{s,1}((a_1,\{j\})|\underset{\sim}{u})$$

$$E\Big[\sum_{i\in a_1} \delta_{i,j,a_1,\{j\}}(\underset{\sim}{u};V_{i_1},\ldots,V_{i_s})\, L(\underset{\sim}{\theta},(a_1,\{j\}),i,j)|\underset{\sim}{U}=\underset{\sim}{u}\Big],$$

and

$D_4(\underset{\sim}{u},(\nu,\eta,\delta))$

$$= \sum_{\substack{s=2\\ s+t\le k}}^{k} \sum_{t=2}^{k-1} \nu_{s,t}(\underset{\sim}{u}) \sum_{\substack{(a_1,a_2):a_1=\{i_1,\ldots,i_s\}\\ a_2=\{j_1,\ldots,j_t\},a_1\cap a_2=\phi}} \eta_{s,t}((a_1,a_2)|\underset{\sim}{u})$$

$$E[\sum_{i\in a_2} \sum_{j\in a_2} \delta_{i,j,a_1,a_2}(\underset{\sim}{u};V_{i_1},\ldots,V_{i_s};V_{j_1},\ldots,V_{j_t})$$

$$L(\underset{\sim}{\theta},(a_1,a_2),i,j)|\underset{\sim}{U}=\underset{\sim}{u}],$$

Now since our loss function satisfies (a) and (b) of (4.1.2), Lemma 4.2 of Eaton (1967) is applicable. Hence

$D_1(\underset{\sim}{u},(\nu,\eta,\delta))$

$$= \nu_{1,1}(\underset{\sim}{u}) \sum_{\substack{i=1\\ i\ne j}}^{k} \sum_{j=1}^{k} \eta_{1,1}((\{i\},\{j\})|\underset{\sim}{u})\ E[L(\underset{\sim}{\theta},(\{i\},\{j\}),i,j)|\underset{\sim}{U}=\underset{\sim}{u}]$$

$$\ge \nu_{1,1}(\underset{\sim}{u}) \sum_{\substack{i=1\\ i\ne j}}^{k} \sum_{j=1}^{k} \eta^*_{1,1}((\{i\},\{j\})|\underset{\sim}{u})\ E[L(\underset{\sim}{\theta},(\{i\},\{j\}),i,j)|\underset{\sim}{U}=\underset{\sim}{u}] \quad \ldots (4.2.4)$$

For $a_1 = \{i_1,\ldots,i_s\}$, $a_2 = \{j_1,\ldots,j_t\}, a_1 \cap a_2 = \phi$, $|a_1|,|a_2| \ge 2$, Writing

$D^{\delta}_{s,t}(\underset{\sim}{u},(a_1,a_2))$

$$= E[\sum_{i\in a_1} \sum_{j\in a_2} \delta_{i,j,a_1,a_2}(\underset{\sim}{u};V_{i_1},\ldots,V_{i_s};V_{j_1},\ldots,V_{j_t})L(\underset{\sim}{\theta},(a_1,a_2),i,j)|\underset{\sim}{U}=\underset{\sim}{u}]$$

we have

$$D_4(\underset{\sim}{u},(\nu,\eta,\delta))$$

$$= \sum_{\substack{s=2 \\ s+t\le k}}^{k} \sum_{t=2}^{k} \nu_{s,t}(\underset{\sim}{u}) \sum_{\substack{(a_1,a_2):a_1=\{i_1,\dots,i_s\} \\ a_2=\{j_1,\dots,j_t\},a_1\cap a_2=\phi}} \eta_{s,t}((a_1,a_2)|\underset{\sim}{u})D^{\delta}_{s,t}(a_1,a_2)) \tag{4.2.5}$$

Now

$$D^{\delta}_{s,t}(\underset{\sim}{u},(a_1,a_2))$$

$$= \int_{\mathbb{R}^{|a_1|+|a_2|}} \sum_{i\in a_1} \sum_{j\in a_2} \delta_{i,j,a_1,a_2}(\underset{\sim}{u};V_{i_1},\dots,V_{i_s};\ V_{j_1},\dots,V_{j_t}) \times \left\{\int_{\mathbb{R}^k} L(\underset{\sim}{\theta},(a_1,a_2),i,j) \prod_{q=1}^{k} f(u_q;\theta_q) \prod_{q\in a_1\cup a_2} h(V_q,\theta_q)d\tau(\underset{\sim}{\theta})\right\} \prod_{q\in a_1\cup a_2} dV_q \times \left[m(\underset{\sim}{u})\right]^{-1}$$

where

$$m(\underset{\sim}{u}) = \int_{\mathbb{R}^k} f(u_q;\theta_q)d\tau(\underset{\sim}{\theta})$$

Using Fubini's Theorem, we get

$$D^{\delta}_{s,t}(\underset{\sim}{u},(a_1,a_2))$$

$$= \int_{\mathbb{R}^k} \sum_{i\in a_1} \sum_{j\in a_2} \delta_{i,j,a_1,a_2}(\underset{\sim}{u};V_{i_1},\dots,V_{i_s};\ V_{j_1},\dots,V_{j_t}) \left\{\int_{\mathbb{R}^k} L(\underset{\sim}{\theta},(a_1,a_2),i,j) \prod_{q=1}^{k} \{f(u_q;\theta_q)h(V_q;\theta_q)\}d\tau(\underset{\sim}{\theta})\right\} \prod_{r=1}^{k} dv_r \times [m(\underset{\sim}{u})]^{-1}$$

$$= \int_{\mathbb{R}^k} \sum_{i\in a_1} \sum_{j\in a_2} \delta_{i,j,a_1,a_2}(\underset{\sim}{u}; V_{i_1}, \ldots, V_{i_s}; V_{j_1}, \ldots, V_{j_t})$$

$$\left\{ \int_{\mathbb{R}^k} L(\underset{\sim}{\theta}, (a_1, a_2), i, j)\, (2\pi n_2)^{-\frac{k}{2}} \exp\left[-\frac{1}{2n_2} \sum_{q=1}^{k} (v_q^2 + \theta_q(u_q + w_q)) \right] \right.$$

$$\left. \exp\left[-\frac{n_1+n_2}{2} \sum_{q=1}^{k} \theta_q^2 \right] d\tau(\underset{\sim}{\theta}) \right\} \prod_{r=1}^{k} dV_r \, [\tilde{m}(\underset{\sim}{u})]^{-1}$$

where

$$\tilde{m}(\underset{\sim}{u}) = \int_{\mathbb{R}^k} \exp\left(\sum_{q=1}^{k} \theta_q (u_q - \frac{n_1}{2} \right) d\tau(\underset{\sim}{\theta}).$$

On substituting $u_q + V_q = w_q$ in $D^{\delta}_{s,t}(\underset{\sim}{u}, (a_1, a_2))$, we get

$D^{\delta}_{s,t}(\underset{\sim}{u}, (a_1, a_2))$

$$= \int_{\mathbb{R}^k} (2\pi n_2)^{-\frac{k}{2}} \exp\left[-\frac{1}{2n_2} \sum_{q=1}^{k} (w_q - u_q)^2 \right]$$

$$\sum_{i\in a_1} \sum_{j\in a_2} \delta_{i,j,a_1,a_2}(\underset{\sim}{u}: w_{i_1} - u_{i_1}, \ldots, w_{i_s} - u_{i_s}; w_{j_1} - u_{j_1}, \ldots, w_{j_t} - u_{j_t})$$

$$\left\{ \int_{\mathbb{R}^k} L(\underset{\sim}{\theta}, (a_1, a_2), i, j) \exp\left[\sum_{q=1}^{k} \theta_q w_q \right] \prod_{r=1}^{k} dw_r \, (\tilde{m}(u))^{-1} \right.$$

where $d\hat{\tau}(\underset{\sim}{\theta}) = \exp\left[-\frac{n_1+n_2}{2} \sum_{q=1}^{k} \theta_q^2 \right] d\tau(\underset{\sim}{\theta}).$

Note that $d\hat{\tau}(\underset{\sim}{\theta})$ is also symmetric in $\underset{\sim}{\theta}$.

For $a_1, a_2 \subset \{1, \ldots, k\}$, $a_1, a_2 \neq \phi$, $a_1 \cap a_2 = \phi$, define

$$E_1((a_1,a_2),i,j|\underset{\sim}{w}) = \int_{\mathbb{R}^k} L(\underset{\sim}{\theta},(a_1,a_2),i,j) \exp\left[\sum_{q=1}^{k} \theta_q w_q\right] d\hat{\tau}(\underset{\sim}{\theta}).$$

Then

$D^{\delta}_{s,t}(\underset{\sim}{u},(a_1,a_2))$

$$= \int_{\mathbb{R}^k} (2\pi n_2)^{-\frac{k}{2}} \exp\left[-\frac{1}{2n_2}\sum_{q=1}^{k}(w_q-u_q)^2\right]$$

$$\sum_{i\in a_1}\sum_{j\in a_2} \delta_{i,j,a_1,a_2}(\underset{\sim}{u};w_{i_1}-u_{i_1},\ldots,w_{i_s}-u_{i_s};w_{j_1}-u_{j_1},\ldots,w_{j_t}-u_{j_t})$$

$$E_1((a_1,a_2),i,j)|\underset{\sim}{w}) \prod_{r=1}^{k} dw_r \left[\tilde{m}(u)\right]^{-1} \qquad (4.2.6)$$

Now we prove the monotonicity of $E_1((a_1,a_2),i,j|\underset{\sim}{w})$.

<u>**Lemma 4.2.1**</u> **:** Suppose $i, i' \in a_1$, and $j, j' \in a_2$, then for every fixed $\underset{\sim}{w} \in \mathbb{R}^k$,

(i) $E_1((a_1,a_2),i,j|\underset{\sim}{w}) \geq E_1((a_1,a_2),i',j|\underset{\sim}{w})$, if $w_i \geq w_{i'}$

(ii) $E_1((a_1,a_2),i,j|\underset{\sim}{w}) \geq E_1((a_1,a_2),i,j'|\underset{\sim}{w})$, if $w_j \geq w_{j'}$

<u>**Proof**</u> :

(i) $E_1((a_1,a_2),i,j|\underset{\sim}{w}) - E_1((a_1,a_2),i',j|\underset{\sim}{w})$

$$= \int_{\{\underset{\sim}{\theta}:\theta_i > \theta_{i'}\}} (L(\underset{\sim}{\theta},(a_1,a_2),i,j)-L(\underset{\sim}{\theta},(a_1,a_2),i',j)) \exp\left[\sum_{q=1}^{k} \theta_q w_q\right] d\hat{\tau}(\underset{\sim}{\theta})$$

$$+ \int_{\{\underset{\sim}{\theta}:\theta_i = \theta_{i'}\}} (L(\underset{\sim}{\theta},(a_1,a_2),i,j)-L(\underset{\sim}{\theta},(a_1,a_2),i',j)) \exp\left[\sum_{q=1}^{k} \theta_q w_q\right] d\hat{\tau}(\underset{\sim}{\theta})$$

$$+ \int_{\{\underset{\sim}{\theta}:\theta_i < \theta_{i'}\}} (L(\underset{\sim}{\theta},(a_1,a_2),i,j)-L(\underset{\sim}{\theta},(a_1,a_2),i',j)) \exp\left[\sum_{q=1}^{k} \theta_q w_q\right] d\hat{\tau}(\underset{\sim}{\theta})$$

Since $\hat{\tau}$ is symmetric, the second integral is zero and the roles of θ_i and $\theta_{i'}$ can be interchanged in 3rd integral so that

$E_1((a_1,a_2),i,j|\underset{\sim}{w}) - E_1((a_1,a_2),i',j|\underset{\sim}{w})$

$$= \int_{\{\underset{\sim}{\theta}:\theta_i > \theta_{i'}\}} (L(\underset{\sim}{\theta},(a_1,a_2),i,j)-L(\underset{\sim}{\theta},(a_1,a_2),i',j)) \prod_{\substack{q=1 \\ q\neq i,i'}}^{k} e^{\theta_q w_q} \left[e^{\theta_i w_i} e^{\theta_{i'} w_{i'}} - e^{\theta_{i'} w_i} e^{\theta_i w_{i'}}\right] d\hat{\tau}(\underset{\sim}{\theta})$$

≥ 0 (on using MLR property and property (c) of (4.1.2))

(ii) $E_1((a_1,a_2),i,j|\underset{\sim}{w}) - E_1((a_1,a_2),i,j'|\underset{\sim}{w})$

$$= \int_{\{\underset{\sim}{\theta}:\theta_{j'}>\theta_j\}} (L(\underset{\sim}{\theta},(a_1,a_2),i,j)-L(\underset{\sim}{\theta},(a_1,a_2),i,j'))\, \exp\left[\sum_{q=1}^{k}\theta_q w_q\right] d\hat{\tau}(\underset{\sim}{\theta})$$

$$+ \int_{\{\underset{\sim}{\theta}:\theta_{j'}=\theta_j\}} (L(\underset{\sim}{\theta},(a_1,a_2),i,j)-L(\underset{\sim}{\theta},(a_1,a_2),i,j'))\, \exp\left[\sum_{q=1}^{k}\theta_q w_q\right] d\hat{\tau}(\underset{\sim}{\theta})$$

$$+ \int_{\{\underset{\sim}{\theta}:\theta_{j'}<\theta_j\}} (L(\underset{\sim}{\theta},(a_1,a_2),i,j)-L(\underset{\sim}{\theta},(a_1,a_2),i,j'))\, \exp\left[\sum_{q=1}^{k}\theta_q w_q\right] d\hat{\tau}(\underset{\sim}{\theta})$$

On using invariance of $\hat{\tau}$ and interchanging the roles of $\theta_{j'}$ and θ_j in 3rd integral, we get

$E_1((a_1,a_2),i,j|\underset{\sim}{w}) - E_1((a_1,a_2),i,j'|\underset{\sim}{w})$

$$= \int_{\{\underset{\sim}{\theta}:\theta_{j'}>\theta_j\}} (L(\underset{\sim}{\theta},(a_1,a_2),i,j) - L(\underset{\sim}{\theta},(a_1,a_2),i,j'))\prod_{\substack{q=1\\ q\neq j,j'}}^{k} e^{\theta_q\omega_q}$$

$$[e^{\theta_j\omega_j}\, e^{\theta_{j'}\omega_{j'}} - e^{\theta_{j'}\omega_j}\, e^{\theta_j\omega_{j'}}]\, d\hat{\tau}(\underset{\sim}{\theta})$$

≥ 0 (on using MLR property and property (d) of (4.1.2). ■

Lemma 4.2.2 : For $a_1 = \{i_1,\dots,i_s\}$, $a_2 = \{j_1,\dots,j_t\}$, $s,t \ge 2$, $a_1 \cap a_2 = \phi$ and for every δ

$$D^{\delta}_{s,t}(\underset{\sim}{u},(a_1,a_2)) \ge D^{\delta^*}_{s,t}(\underset{\sim}{u},(a_1,a_2)) \qquad (4.2.7)$$

where δ^* is same as defined by (4.2.1).

Proof : On using Lemma 4.2.1, it follows that for every fixed $\underset{\sim}{w} \in \mathbb{R}^k$

$$\sum_{i\in a_1}\sum_{j\in a_2} \delta_{i,j,a_1,a_2}(\underset{\sim}{u};w_{i_1}-u_{i_1},\dots,w_{i_s}-u_{i_s};w_{j_1}-u_{j_1},\dots,w_{j_t}-u_{j_t}) \, E_1((a_1,a_2),i,j|\underset{\sim}{w})$$

$$\ge \sum_{i\in a_1}\sum_{j\in a_2} \delta^*_{i,j,a_1,a_2}(\underset{\sim}{u};w_{i_1}-u_{i_1},\dots,w_{i_s}-u_{i_s};\ w_{j_1}-u_{j_1},\dots,w_{j_t}-u_{j_t}) \, E_1((a_1,a_2),i,j|\underset{\sim}{w})$$

Now the result follows from (4.2.6). ■

For $i,j = 1,\dots,k$, $a_1 = \{i_1,\dots,i_s\}$, $a_2 = \{j_1,\dots,j_t\}$, $2 \le s \le k-1$, $2 \le t \le k-1$, $i \notin a_2$, $j \notin a_1$, let

$$D^{\delta}_{1,t}(\underset{\sim}{u},(\{i\},a_2))$$
$$= E\Big[\sum_{j\in a_2} \delta_{i,j,\{i\},a_2}(\underset{\sim}{u};V_{j_1},\dots,V_{j_t})\ L(\underset{\sim}{\theta},(\{i\},a_2),i,j)\,\Big|\,\underset{\sim}{U}=\underset{\sim}{u}\Big]$$

and

$$D^{\delta}_{s,1}(\underset{\sim}{u},(a_1,\{j\}))$$
$$= E\Big[\sum_{i\in a_1} \delta_{i,j,a_1,\{j\}}(\underset{\sim}{u};V_{i_1},\dots,V_{i_s})\ L(\underset{\sim}{\theta},(a_1,\{j\}),i,j)\,\Big|\,\underset{\sim}{U}=\underset{\sim}{u}\Big]$$

then

$D_2(\underset{\sim}{u},(\nu,\eta,\delta))$

$$= \sum_{i=1}^{k} \sum_{t=2}^{k-1} \nu_{1,t}(\underset{\sim}{u}) \sum_{\substack{a_2:a_2=\{j_1,\ldots,j_t\} \\ j_q\neq i,q=1,\ldots,t}} \eta_{1,t}((\{i\},a_2)|\underset{\sim}{u})$$

$$D^{\delta}_{1,t}(\underset{\sim}{u},(\{i\},a_2)) \qquad (4.2.8)$$

and

$D_3(\underset{\sim}{u},(\nu,\eta,\delta))$

$$= \sum_{s=2}^{k-1} \sum_{j=1}^{k} \nu_{s,1}(\underset{\sim}{u}) \sum_{\substack{a_1:a_1=\{i_1,\ldots,i_s\} \\ i_q\neq j,q=1,\ldots,s}} \eta_{s,1}((a_1,\{j\})|\underset{\sim}{u})$$

$$D^{\delta}_{s,1}(\underset{\sim}{u},(a_1,\{j\})) \qquad (4.2.9)$$

Now,

$$D^{\delta}_{1,t}(\underset{\sim}{u},(\{i\},a_2)) = \int_{\mathbb{R}^k} (2\pi n_2)^{-\frac{k}{2}} \exp\left(-\frac{1}{2n_2}\sum_{q=1}^{k}(w_q-u_q)^2\right)$$

$$\sum_{j\in a_2} \delta_{i,j,\{i\},a_2}(\underset{\sim}{u};w_{j_1}-u_{j_1},\ldots,w_{j_t}-u_{j_t})$$

$$E_1((\{i\},a_2),i,j|\underset{\sim}{w}) \prod_{r=1}^{k} dw_r\left[\tilde{m}(\underset{\sim}{u})\right]^{-1}$$

and

$$D^{\delta}_{s,1}(\underset{\sim}{u},(a_1,\{j\})) = \int_{\mathbb{R}^k} (2\pi n_2)^{-\frac{k}{2}} \exp\left[-\frac{1}{2n_2}\sum_{q=1}^{k}(w_q-u_q)^2\right]$$

$$\sum_{i\in a_1} \delta_{i,j,a_1,\{j\}}(\underset{\sim}{u};w_{i_1}-u_{i_1},\ldots,w_{i_s}-u_{i_s})$$

$$E_1((a_1,\{j\}),i,j|\underset{\sim}{\omega}) \prod_{i=1}^{k} dw_r\left[\tilde{m}(\underset{\sim}{u})\right]^{-1}$$

The next lemma is an immediate consequence of Lemma 4.2.1.

Lemma 4.2.3 : For $i,i' \in a_1$, $j,j' \in a_2$, $i,i' \notin a_2$ and $j,j' \notin a_1$,

(i) $D^{\delta}_{s,1}(\underset{\sim}{u},(a_1,\{j\})) \geq D^{\delta^*}_{s,1}(\underset{\sim}{u},(a_1,\{j\}))$, $2 \leq s \leq k-1$, (4.2.10)

(ii) $D^{\delta}_{1,t}(\underset{\sim}{u},(\{i\},a_2)) \geq D^{\delta^*}_{1,t}(\underset{\sim}{u},(\{i\},a_2))$, $2 \leq t \leq k-1$, (4.2.11)

The following theorem proves the optimality of the decision rule δ^*.

Theorem 4.2.1 : Let τ be a symmetric prior and suppose that loss function L satisfies (4.1.1) and (4.1.2), then

(i) $r(\tau,(\nu,\tilde{\eta},\delta^*)) \leq r(\tau,(\nu,\eta,\delta))$, $\forall\ (\nu,\eta,\delta) \in \mathfrak{D}$

and

(ii) $R(\underset{\sim}{\theta},(\nu,\tilde{\eta},\delta^*)) \leq R(\underset{\sim}{\theta},(\nu,\eta,\delta))$, $\forall\ (\nu,\eta,\delta) \in \mathfrak{D}_I$

where $\tilde{\eta}$ is same as η except $\tilde{\eta}_{1,1} = \eta^*_{1,1}$.

Proof :

(i) From (4.2.4) and from the definition of $\tilde{\eta}$, it follows that

$D_1(\underset{\sim}{u},(\nu,\eta,\delta)) \geq D_1(\underset{\sim}{u},(\nu,\tilde{\eta},\delta))$, $\forall\ (\nu,\eta,\delta) \in \mathfrak{D}$ (4.2.12)

Also from (4.2.5) and (4.2.7), we get that

$$D_4(\underset{\sim}{u},(\nu,\eta,\delta)) \geq D_4(\underset{\sim}{u},(\nu,\eta,\delta^*)), \quad \forall\ (\nu,\eta,\delta) \in \mathfrak{D} \tag{4.2.13}$$

Similarly,

$$D_2(\underset{\sim}{u},(\nu,\eta,\delta)) \geq D_2(\underset{\sim}{u},(\nu,\eta,\delta^*)), \quad \forall\ (\nu,\eta,\delta) \in \mathfrak{D}, \tag{4.2.14}$$

(from (4.2.8) and (4.2.11))

and

$$D_3(\underset{\sim}{u},(\nu,\eta,\delta)) \geq D_3(\underset{\sim}{u},(\nu,\eta,\delta^*)), \quad \forall\ (\nu,\eta,\delta) \in \mathfrak{D} \tag{4.2.15}$$

(from (4.2.9) and (4.2.10)

Now on using (4.2.12) - (4.2.15) in (4.2.3), we get

$$r^*_{\underset{\sim}{u}}((\nu,\eta,\delta)) \geq r^*_{\underset{\sim}{u}}((\eta,\tilde{\eta},\delta^*))$$

which proves the first assertion.

(ii) For $(\nu,\eta,\delta) \in \mathfrak{D}_I$

$$R(\underset{\sim}{\theta},(\nu,\eta,\delta)) = R(g\underset{\sim}{\theta},(\nu,\eta,\delta)), \quad \forall\ g \in G \tag{4.2.16}$$

Fix $\underset{\sim}{\theta} \in \mathbb{R}^k$ and let τ be a prior which puts equal weight on each of permutation of $\underset{\sim}{\theta}$. Then on using (i), we get

$$\frac{1}{\underline{|k}} \sum_{g \in G} R(g\underset{\sim}{\theta},(\nu,\tilde{\eta},\delta^*)) \leq \frac{1}{\underline{|k}} \sum_{g \in G} R(g\underset{\sim}{\theta},(\nu,\eta,\delta))$$

$$\Longrightarrow R(\underset{\sim}{\theta},(\nu,\tilde{\eta},\delta^*)) \leq R(\underset{\sim}{\theta},(\nu,\eta,\delta)),\ \forall\ (\nu,\eta,\delta) \in \mathfrak{D}_I$$

(from (4.2.16)). ∎

4.3 The Optimal Two-stage Decision rule

Now for $1 \le s,t \le k-1$, $3 \le s+t \le k$, we can write

$$D^{\delta^*}_{s,t}(\underset{\sim}{u},(a_1,a_2))$$

$$= \int_{\mathbb{R}^k} \left\{ \underset{i\in a_1,\, j\in a_2}{\mathrm{Min}} L(\underset{\sim}{\theta},(a_1,a_2),i,j) \exp\left[\sum_{q=1}^{k} \theta_q w_q\right] d\hat{\tau}(\underset{\sim}{\theta}) \right\}$$

$$(2\pi n_2)^{-\frac{k}{2}} \exp\left[-\frac{1}{2n_2} \sum_{q=1}^{k} (w_q-u_q)^2\right] \prod_{r=1}^{k} dw_r \quad [\tilde{m}(u)]^{-1}$$

Define

$$L^*(\underset{\sim}{w},(a_1,a_2)) = \underset{i\in a_1,\, j\in a_2}{\min} \int_{\mathbb{R}^k} L(\underset{\sim}{\theta},(a_1,a_2),i,j) \exp\left[\sum_{q=1}^{k} \theta_q w_q\right] d\hat{\tau}(\underset{\sim}{\theta})$$

$$= \underset{i\in a_1,\, j\in a_2}{\min} E_1((a_1,a_2),i,j|\underset{\sim}{w}).$$

The following lemma proves the monotonicity of L^*.

Lemma 4.3.1: Let $a_1 = \{i_1,\dots,i_{r-1},i\}$, $a_2 = \{j_1,\dots,j_{s-1};j\}$, $\tilde{a}_1 = \{i_1,\dots,i_{r-1},i'\}$, $\tilde{a}_2 = \{j_1,\dots,j_{s-1},j'\}$, $|a_1|,|a_2|,|\tilde{a}_1|,|\tilde{a}_2| \ge 1$, $r+s \ge 3$, with $a_1 \cap a_2 = \phi$, $\tilde{a}_1 \cap \tilde{a}_2 = \phi$, $\tilde{a}_1 \cap a_2 = \phi$, $a_1 \cap \tilde{a}_2 = \phi$, then

(i) $L^*(\underset{\sim}{w},(a_1,a_2)) \le L^*(\underset{\sim}{w},(\tilde{a}_1,a_2))$, if $w_{i'} \ge w_i$

(ii) $L^*(\underset{\sim}{w},(a_1,a_2)) \le L^*(\underset{\sim}{w},(a_1,\tilde{a}_2))$, if $w_j \ge w_{j'}$.

Proof :

(i) For $i \in a_1$, $j \in a_2$, $a_1 \cap a_2 = \phi$

$$E_1((\tilde{a}_1,a_2),q,m|\underset{\sim}{w}) - E_1((a_1,a_2),q,m|\underset{\sim}{w})$$

$$= \int_{\{\underset{\sim}{\theta}:\theta_{i'}>\theta_i\}} [L(\underset{\sim}{\theta},(\tilde{a}_1,a_2),q,m)-L(\underset{\sim}{\theta},(a_1,a_2),q,m)] \exp\left(\sum_{q=1}^{k}\theta_q w_q\right) d\hat{\tau}(\underset{\sim}{\theta})$$

$$+ \int_{\{\underset{\sim}{\theta}:\theta_{i'}=\theta_i\}} [L(\underset{\sim}{\theta},(\tilde{a}_1,a_2),q,m)-L(\underset{\sim}{\theta},(a_1,a_2),q,m)] \exp\left(\sum_{q=1}^{k}\theta_q w_q\right) d\hat{\tau}(\underset{\sim}{\theta})$$

$$+ \int_{\{\underset{\sim}{\theta}:\theta_{i'}<\theta_i\}} [L(\underset{\sim}{\theta},(\tilde{a}_1,a_2),q,m)-L(\underset{\sim}{\theta},(a_1,a_2),q,m)] \exp\left(\sum_{q=1}^{k}\theta_q w_q\right) d\hat{\tau}(\underset{\sim}{\theta})$$

Since $\hat{\tau}$ is symmetric, the second integral is zero and the roles of θ_i and $\theta_{i'}$ can be interchanged in the third integral so that

$$E_1((\tilde{a}_1,a_2),q,m|\underset{\sim}{w}) - E_1((a_1,a_2),q,m|\underset{\sim}{w})$$

$$= \int_{\{\underset{\sim}{\theta}:\theta_{i'}>\theta_i\}} [L(\underset{\sim}{\theta},(\tilde{a}_1,a_2),q,m)-L(\underset{\sim}{\theta},(a_1,a_2),q,m)] \prod_{\substack{q=1\\ q\neq i,i'}}^{k} e^{\theta_q w_q}$$

$$[e^{\theta_i w_i}\, e^{\theta_{i'} w_{i'}} - e^{\theta_i w_{i'}}\, e^{\theta_{i'} w_i}]\, d\hat{\tau}(\underset{\sim}{\theta})$$

≥ 0, if $w_{i'} \geq w_i$

(using MLR property and property ((e) (i)) of (4.1.2))

Again using the same argument, we get

$$E_1((\tilde{a}_1,a_2),i',m|\underset{\sim}{w}) - E_1((a_1,a_2),i,m|\underset{\sim}{w}) \geq 0$$

(ii) Follows in a similar manner. ∎

Now we present the main theorem of this chapter.

Theorem 4.3.1 : Suppose τ be a symmetric prior and suppose loss function L satisfies (4.1.1) and (4.1.2). Then

(i) $r(\tau,(\nu,\eta^*,\delta^*)) \le r(\tau,(\nu,\eta,\delta)), \quad \forall\ (\nu,\eta,\delta) \in \mathfrak{D}$

and

(ii) $R(\underset{\sim}{\theta},(\nu,\eta^*,\delta^*)) \le R(\underset{\sim}{\theta},(\nu,\eta,\delta)), \quad \forall\ (\nu,\eta,\delta) \in \mathfrak{D}_I$

Proof : From Theorem 4.2.1, we have

$$r(\tau,(\nu,\tilde{\eta},\delta^*)) \le r(\tau,(\nu,\eta,\delta)), \quad \forall\ (\nu,\eta,\delta) \in \mathfrak{D} \tag{4.3.1}$$

For s,t ≥ 1, 3 ≤ s+t ≤ k, consider the problem of partitioning {1,...,k} into three (two) disjoint subsets of sizes s, k-s-t, and t (s and t) with s+t < k (s+t = k). Using Lemma 4.2.5, it is easy to see that $L^*(\underset{\sim}{\omega},(a_1,a_2))$ has the property of the loss function assumed by Eaton (1967). Hence for 1 ≤ s, t ≤ k, 3 ≤ s+t ≤ k,

$$\sum_{\substack{(a_1,a_2):|a_1|=s,\\ |a_2|=t,a_1\cap a_2=\phi}} \eta_{s,t}((a_1,a_2)|\underset{\sim}{u})\ D^{\delta^*}_{s,t}(\underset{\sim}{u},(a_1,a_2))$$

$$= \sum_{\substack{(a_1,a_2):|a_1|=s,\\ |a_2|=t,a_1\cap a_2=\phi}} \eta_{s,t}((a_1,a_2)|\underset{\sim}{u})$$

$$\int_{\mathbb{R}^k} L^*(\underset{\sim}{\omega},(a_1,a_2))\ (2\pi n_2)^{-\frac{k}{2}} \exp\left[-\frac{1}{2n_2}\sum_{q=1}^{k}(w_q-u_q)^2\right]$$

$$\prod_{r=1}^{k} d\omega_r\ [\tilde{m}(\underset{\sim}{u})]^{-1}$$

$$\geq \sum_{\substack{(a_1,a_2):|a_1|=s,\\ |a_2|=t,a_1\cap a_2=\phi}} \eta^*_{s,t}((a_1,a_2)|\underset{\sim}{u})$$

$$\int_{\mathbb{R}^k} L^*(\underset{\sim}{\omega},(a_1,a_2))\,(2\pi n_2)^{-\frac{k}{2}} \exp\left[-\frac{1}{2n_2}\sum_{q=1}^{k}(w_q-u_q)^2\right]$$

$$\prod_{r=1}^{k} d\omega_r \,[\tilde{m}(\underset{\sim}{u})]^{-1}$$

$$= \sum_{\substack{(a_1,a_2):|a_1|=s,\\ |a_2|=t,a_1\cap a_2=\phi}} \eta^*_{s,t}((a_1,a_2)|\underset{\sim}{u})D^{\delta^*}_{s,t}(\underset{\sim}{u},(a_1,a_2)) \qquad (4.3.2)$$

Now the result follows from (4.2.3), (4.3.1) and (4.3.2). ■

Corollary 4.3.1 : Under the assumptions of Theorem 4.2.1, if for a given η (ν,η^*,δ^*) is minimax in $\mathcal{D}_I$, then (ν,η^*,δ^*) is minimax in $\mathcal{D}$.

Proof. Since the group of permutations G is finite, the result follows from Blackwell and Girshick (1954, Chapter ~~8~~ pp. 223-229). ■

We conclude this chapter with the following remark.

Remark. Let the observations from π_1 have a density $C(\theta_1)$ $\exp(\theta_1 x)$ $b(x)$ with respect to Lebesgue measure or counting measure, then as in Gupta and Miescke (1983), all the results of this chapter can be generalized to the case where the underlying density is strongly unimodal.

CHAPTER V

SELECTING THE BETTER OF TWO BINOMIAL POPULATIONS : OPTIMAL DECISION RULES

5.1 Introduction :

Let $\pi_1,\ldots,\pi_k$ be $k(\geq 2)$ independent binomial populations with single trial success probabilities $\theta_1,\ldots,\theta_k$ respectively. Let n_i be the number of independent observations from π_i, $i = 1,\ldots,k$. The goal is to select the UEP (the population associated with and $\theta_{[k]}$).

For selecting the UEP, Sobel and Huyett (1957) considered the problem of finding a smallest common sample size such that the probability of CS is more than or equal to a specified P^* $(\frac{1}{k} < P^* < 1)$ for a specified indifference-zone. Later Hall (1959) proved that the rule proposed by Sobel and Huyett is minimax for equal sample sizes. For the case of unequal sample sizes Gupta and Sobel (1958, 1960) suggest a decision rule which selects the population corresponding to largest sample proportion of successes as the population associated with $\theta_{[k]}$. We call the decision rule proposed by Gupta and Sobel as intuitive decision rule.

Note that for $k = 2$, the goal of simultaneously selecting the LEP and the UEP is equivalent to the goal of selecting UEP.

Risko (1985) considered the problem of finding a minimax rule for the case of $k = 2$ binomial populations with unequal sample sizes. It was observed that intuitive decision rule performs very

badly when one sample size is very large compared to the other. For this case Risko suggested a decision rule which is minimax when one sample size goes to infinity with other sample size remaining fixed and finite. For the case where both the decision rules result in $P_{\underset{\sim}{\theta}}(CS)$ less than $\frac{1}{2}$ a class of decision rules which depends on two parameters is given and an attempt is made to find a minimax rule in this class. It is observed that this restricted minimax rule possesses many properties consistent with globally minimax rule and that it is in fact globally minimax for certain configurations of sample sizes.

Following Risko, we consider the problem of characterizing minimax rules for selecting the better of two binomial populations with unequal sample sizes. We derive some necessary conditions for a rule to be globally minimax. Several numerical examples are also considered.

5.2 Formulation of the Problem :

Let π_1 and π_2 be two independent binomial populations with single trial success probabilities θ_1 and θ_2 and sample sizes n_1 and n_2 respectively. Suppose X_1 and X_2 denote independent random variables representing π_1 and π_2 respectively. Our aim is to select the population associated with larger of θ_1 and θ_2.

Here, a decision rule δ is a map from $(0,1,\ldots,n_1) \times (0,1,\ldots,n_2)$ to $[0,1]$ with the following interpretation :

If $X_1 = i$ and $X_2 = j$ is observed then $\delta(i,j) \in [0,1]$ is the probability of selecting π_1 and $1-\delta(i,j)$ is the probability of selecting π_2.

Let $\mathcal{D}^{n_1,n_2}$ denote the class of all decision rules for the problem at hand and let

$$\mathcal{D}_1^{n_1,n_2} = \{\delta \in \mathcal{D}^{n_1,n_2} : \delta(i,j) = 1-\delta(n_1-i,\ n_2-j)\}$$

Risko (1985) proved that class $\mathcal{D}_1^{n_1,n_2}$ is an essentially complete class for finding a minimax rule. Let

$$B_{n_1,i}(\theta_1) = \binom{n_1}{i} \theta_1^i (1-\theta_1)^{n_1-i}, \quad i = 0,1,\ldots,n_1$$

$$B_{n_2,j}(\theta_2) = \binom{n_2}{j} \theta_2^j (1-\theta_2)^{n_2-j}, \quad j = 0,1,\ldots,n_2$$

Let $P(CS|\delta)$ denote probability of CS using rule δ and suppose

$$\begin{aligned} Q(\delta) &= P_{\underset{\sim}{\theta}} \text{ (selecting } \pi_1 \text{ using rule } \delta) \\ &= \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} \delta(i,j)\, B_{n_1,i}(\theta_1)\, B_{n_2,j}(\theta_2) \end{aligned}$$

where $\underset{\sim}{\theta} = (\theta_1, \theta_2)$.

Let

$$\begin{aligned} \overline{Q}(\delta) &= \min_{|\theta_1-\theta_2|\geq d} P_{\underset{\sim}{\theta}}(CS|\delta) \\ &= \min\left\{ \min_{\theta_1\geq\theta_2+d} Q(\delta),\ \min_{\theta_2\geq\theta_1+d} (1-Q(\delta))\right\} \end{aligned}$$

where d $(0 < d < 1)$ is pre-specified.

We desire to find a minimax rule in $\mathcal{D}^{n_1,n_2}$, that is, a $\delta_0 \in \mathcal{D}^{n_1,n_2}$ such that

$$\overline{Q}(\delta_o) = \max_{\delta \in \mathcal{D}^{n_1,n_2}} \overline{Q}(\delta)$$

In the next section, we give some necessary conditions for a rule to be minimax.

5.3 Some Necessary Conditions For Minimaxity

Let $\mathcal{D}_o^{n_1,n_2} \subset \mathcal{D}^{n_1,n_2}$ consist of the δ_o's for which $\overline{Q}(\delta_o) = \max_{\delta \in \mathcal{D}^{n_1,n_2}} \overline{Q}(d)$, that is $\mathcal{D}_o^{n_1,n_2}$ is the class of minimax decision rules. Then we have the following lemmas :

Lemma 5.3.1 : $\mathcal{D}_o^{n_1,n_2}$ is a convex set.

Proof : Let $\delta_o^o, \delta_o' \in \mathcal{D}_o^{n_1,n_2}$ and $0 \le \alpha \le 1$.

Consider $\delta_o^\alpha = \alpha\, \delta_o^o + (1-\alpha)\, \delta_o'$. Clearly $\delta_o^\alpha \in \mathcal{D}^{n_1,n_2}$ and

$$Q(\delta_o^\alpha) = \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} (\alpha\, \delta_o^o(i,j) + (1-\alpha)\, \delta_o'(i,j))\, B_{n_1,i}(\theta_1)\, B_{n_2,j}(\theta_2)$$

$$= \alpha \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} \delta_o^o(i,j)\, B_{n_1,i}(\theta_1)\, B_{n_2,j}(\theta_2)$$

$$+ (1-\alpha) \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} (\delta_o'(i,j)\, B_{n_1,i}(\theta_1)\, B_{n_2,j}(\theta_2)$$

$$= \alpha\, Q(\delta_o^o) + (1-\alpha)\, Q(\delta_o')$$

$$1-Q(\delta_o^\alpha) = 1-\alpha\, Q(\delta_o^o) - (1-\alpha)\, Q(\delta_o')$$

$$= \alpha(1-Q(\delta_o^o)) + (1-\alpha)(1-Q(\delta_o'))$$

$$\overline{Q}(\delta_o^\alpha) = \min\left\{\min_{\theta_1\ge\theta_2+d} Q(\delta_o^\alpha)\ ,\ \min_{\theta_2\ge\theta_1+d} (1-Q(\delta_o^\alpha))\right\}$$

$$= \min\left\{\min_{\theta_1\ge\theta_2+d} (\alpha\, Q(\delta_o^0) + (1-\alpha)\, Q(\delta_o')),\right.$$

$$\left.\min_{\theta_2\ge\theta_1+d} (\alpha(1-Q(\delta_o^0)) + (1-\alpha)(1-Q(\delta_o')))\right\}$$

Since $\min\limits_{x\in B} (f(x) + g(x)) \ge \min\limits_{x\in B} f(x) + \min\limits_{x\in B} g(x)$ for any set B, we have

$$\overline{Q}(\delta_o^\alpha) \ge \min\left\{\alpha \min_{\theta_1\ge\theta_2+d} Q(\delta_o^0) + (1-\alpha) \min_{\theta_1\ge\theta_2+d} Q(\delta_o'),\right.$$

$$\left.\alpha \min_{\theta_2\ge\theta_1+d} (1-Q(\delta_o^0)) + (1-\alpha) \min_{\theta_2\ge\theta_1+d} (1-Q(\delta_o'))\right\}$$

$$\ge \alpha\, \overline{Q}(\delta_o^0) + (1-\alpha)\, \overline{Q}(\delta_o')$$

$$= \max_{\delta\in\mathcal{D}^{n_1,n_2}} \overline{Q}(\delta).$$

Hence $\overline{Q}(\delta_o^\alpha) = \max\limits_{\delta\in\mathcal{D}^{n_1,n_2}} \overline{Q}(\delta)$, which implies $\delta_o^\alpha \in \mathcal{D}_o^{n_1,n_2}$, $\forall\ \alpha\in[0,1]$. ■

Lemma 5.3.2 : Let $\delta \in \mathcal{D}^{n_1,n_2}$ and define $\tilde{\delta}$ by

$$\tilde{\delta}(i,j) = 1-\delta(n_1-i,\ n_2-j),\ i = 0,1,\ldots,n_1,\ j = 0,1,\ldots,n_2.$$

Then $\delta \in \mathcal{D}_o^{n_1,n_2}$ if and only if $\tilde{\delta} \in \mathcal{D}_o^{n_1,n_2}$.

Proof : $$Q(\delta) = \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} \delta(i,j)\, B_{n_1,i}(\theta_1)\, B_{n_2,j}(\theta_2)$$

$$= \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} \delta(i,j)\binom{n_1}{i}\binom{n_2}{j}\theta_1^i(1-\theta_1)^{n_1-i}\,\theta_2^j(1-\theta_2)^{n_2-j}$$

$$= \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} \delta(i,j) \begin{pmatrix} n_1 \\ n_1 - i \end{pmatrix} \begin{pmatrix} n_2 \\ n_2 - j \end{pmatrix} \theta_1^i \, (1-\theta_1)^{n_1 - i}$$

$$\theta_2^j \, (1-\theta_2)^{n_2 - j}$$

$$= \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} \delta(n_1 - i, n_2 - j) \begin{pmatrix} n_1 \\ i \end{pmatrix} \begin{pmatrix} n_2 \\ j \end{pmatrix} (1-\theta_1)^i \, \theta_1^{n_1 - i}$$

$$(1-\theta_2)^j \quad \theta_2^{n_2 - k}$$

$$= \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} \delta(n_i - i, n_2 - j) \; B_{n_1, i} \, (1-\theta_1) \quad B_{n_2, j}(1-\theta_2)$$

$$= \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} (1-\tilde{\delta}(n_1 - i, n_2 - j)) \; B_{n_1, i}(1-\theta_1) \; B_{n_2, j}(1-\theta_2)$$

$$\min_{\theta_1 \geq \theta_2 + d} Q(\delta) = \min_{1-\theta_2 \geq 1-\theta_1 + d} Q(\delta)$$

$$= \min_{1-\theta_2 \geq 1-\theta_1 + d} \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} (1-\tilde{\delta}(i,j)) \; B_{n_1, i}(1-\theta_1)$$

$$B_{n_2, j}(1-\theta_2)$$

$$= \min_{\theta_2 \geq \theta_1 + d} \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} (1-\tilde{\delta}(i,j)) \; B_{n_1, i}(\theta_1) \; B_{n_2, j}(\theta_2)$$

$$= \min_{\theta_2 \geq \theta_1 + d} (1-Q(\tilde{\delta}))$$

Similarly,

$$\min_{\theta_2 \geq \theta_1 + d} (1-Q(\delta)) = \min_{\theta_1 \geq \theta_2 + d} Q(\tilde{\delta}).$$

Hence

$$\overline{Q}(\delta) = \overline{Q}(\tilde{\delta})$$

and the result follows. ■

As an immediate consequence of above two lemmas we have :

Corollary 5.3.1 : If $\delta_o \in \mathcal{D}_o^{n_1,n_2}$, then

$$\delta_o^{\alpha}(i,j) = \alpha\,\delta_o(i,j) + (1-\alpha)(1-\delta_o(n_1-i,n_2-j)) \in \mathcal{D}_o^{n_1,n_2}$$

$$\forall\ \alpha \in [0,1]$$

Remark : Corollary 5.3.1 suggest that if minimax decision rule δ_o is unique then $\delta_o \in \mathcal{D}_1^{n_1,n_2}$.

Now, $\delta \in \mathcal{D}^{n_1,n_2}$ can be represented by a matrix of order $(n_1+1) \times (n_2+1)$ as follows :

$$\delta \equiv (\delta(i,j)),\ i = 0,1,\ldots,n_1,\ j = 0,1,\ldots,n_2$$

Let

$$\delta^T \equiv (\delta(j,i))\ ,\ i = 0,1,\ldots,n_1,\ j = 0,1,\ldots,n_2,$$

$$1-\delta \equiv (1-\delta(i,j)),\ i = 0,1,\ldots,n_1,\ j = 0,1,\ldots,n_2,$$

$$\approx\!\!\!\!\delta \equiv 1-\delta^T.$$

Lemma 5.3.3 : Let $\delta_o \in \mathcal{D}^{n_1,n_2}$. Then $\delta_o \in \mathcal{D}_o^{n_1,n_2}$ if and only if $\overset{\approx}{\delta}_o \in \mathcal{D}_o^{n_2,n_1}$.

Proof : Let $\delta_o \in \mathcal{D}^{n_1,n_2}$. Then

$$\min_{\theta_1 \geq \theta_2 + d} Q_{n_1,n_2}(\delta_o)$$

$$= \min_{\theta_1 \geq \theta_2 + d} \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} \delta_o(i,j)\ B_{n_1,i}(\theta_1)\ B_{n_2,j}(\theta_2)$$

Here n_1, n_2 is indicated with Q for clarity.

$$\min_{\theta_1 \geq \theta_2 + d} Q_{n_1,n_2}(\delta_o)$$

$$= \min_{\theta_1 \geq \theta_2 + d} \left\{ 1 - \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} (1-\delta_o(i,j))\ B_{n_1,i}(\theta_1)\ B_{n_2,j}(\theta_2) \right\}$$

$$= \min_{\theta_2 \geq \theta_1 + d} \left\{ 1 - \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} (1-\delta_o(i,j))\ B_{n_1,i}(\theta_2)\ B_{n_2,j}(\theta_1) \right\}$$

$$= \min_{\theta_2 \geq \theta_1 + d} \left\{ 1 - \sum_{j=0}^{n_1} \sum_{i=0}^{n_2} (1-\delta_o(j,i))\ B_{n_1,j}(\theta_2)\ B_{n_2,i}(\theta_1) \right\}$$

$$= \min_{\theta_2 \geq \theta_1 + d} \left\{ 1 - \sum_{j=0}^{n_1} \sum_{i=0}^{n_2} \overset{\approx}{\delta}_o(i,j))\ B_{n_2,i}(\theta_1)\ B_{n_1,j}(\theta_2) \right\}$$

$$= \min_{\theta_2 \geq \theta_1 + d} \left\{ 1 - Q_{n_2,n_1}(\overset{\approx}{\delta}_o) \right\}$$

Similarly,

$$\min_{\theta_2 \geq \theta_1 + d} \left[1 - Q_{n_1,n_2}(\delta_o) \right] = \min_{\theta_1 \geq \theta_2 + d} Q_{n_2,n_1}(\overset{\approx}{\delta}_o)$$

$$\Rightarrow \overline{Q}_{n_1,n_2}(\delta_o) = \overline{Q}_{n_2,n_1}(\overset{\approx}{\delta})$$

Now using the fact that $\delta \in \mathcal{D}^{n_1,n_2}$ if and only if $\overset{\approx}{\delta} \in \mathcal{D}^{n_2,n_1}$ we get the required result.■

Following Theorem gives three necessary conditions for a rule to be minimax.

<u>**Theorem 5.3.1**</u> **:** Let $\delta_o \in \mathcal{D}_o^{n_1,n_2}$. Then

(i) $\displaystyle \min_{\theta_1 \geq \theta_2 + d} Q(\delta_o) = \min_{\theta_2 \geq \theta_1 + d} (1-Q(\delta_o))$

(ii) $\delta_o(i_o,j_o) > 0$ for some (i_o,j_o) implies that $\delta_o(i_o+1,j_o-1)=1$

(iii) For $d > .5$, $\delta_o(i_o,j_o) > 0$ for some (i_o,j_o) implies in addition to (ii), that $\delta_o(i_o+1,j_o) = 1$ and $\delta_o(i_o,j_o-1) = 1$.

Proof : (i) Let $M = \min\limits_{\theta_1 \geq \theta_2 + d} Q(\delta_o)$ and $M' = \min\limits_{\theta_2 \geq \theta_1 + d} (1-Q(\delta_o))$.

First suppose that $M' < M$.

Let $\frac{M'}{M} < c < 1$ and consider the decision rule δ_c defined by :

$$\delta_c(i,j) = c\,\delta_o(i,j),\quad i = 0,1,\ldots,n_1,\quad j = 0,1,\ldots,n_2.$$

Then

$$\min_{\theta_2 \geq \theta_1 + d} (1-Q(\delta_c)) = \min_{\theta_2 \geq \theta_1 + d} (1-c\sum_{i=0}^{n_1}\sum_{j=0}^{n_2} \delta_o(i,j)B_{n_1,i}(\theta_1)B_{n_2,j}(\theta_2))$$

$$= \min_{\theta_2 \geq \theta_1 + d} (1-c\sum_{i=0}^{n_1}\sum_{j=0}^{n_2} (-1+\delta_o(i,j)+1)$$

$$B_{n_1,i}(\theta_1)B_{n_2,j}(\theta_2))$$

$$= \min_{\theta_2 \geq \theta_1 + d} (1-c + c\sum_{i=0}^{n_1}\sum_{j=0}^{n_2} (1-\delta_o(i,j))$$

$$B_{n_1,i}(\theta_1)B_{n_2,j}(\theta_2))$$

$$= 1-c+c\min_{\theta_2 \geq \theta_1 + d} \sum_{i=0}^{n_1}\sum_{j=0}^{n_2} (1-\delta_o(i,j))$$

$$B_{n_1,i}(\theta_1)B_{n_2,j}(\theta_2))$$

$$= 1-c+cM'$$

$$> M' \qquad \text{(since } M' < 1\text{).}$$

and

$$\min_{\theta_1 \geq \theta_2 + d} Q(\delta_c) = cM > M',$$

contradicting the assumption that $\delta_0 \in \mathcal{D}_0^{n_1, n_2}$.

Next, suppose $M < M'$.

Let $\frac{M}{M'} < \gamma < 1$, and consider the decision rule δ^{γ} defined by

$\delta^{\gamma}(i,j) = \gamma\, \delta_0(i,j) + (1-\gamma)\, \delta^{u}(i,j)$, $i = 0,1,\ldots,n_1$, $j = 0,1,\ldots,n_2$

where $\delta^{u}(i,j) = 1$, $\forall\ i,j$. Then

$$\min_{\theta_2 \geq \theta_1 + d} (1 - Q(\delta^{\gamma}))$$

$$= \min_{\theta_2 \geq \theta_1 + d} \Big(1 - \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} (\gamma\, \delta_0(i,j) + (1-\gamma)) B_{n_1,i}(\theta_1) B_{n_2,j}(\theta_2)\Big)$$

$$= \min_{\theta_2 \geq \theta_1 + d} \Big(1 - \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} (-\gamma\, (1 - \delta_0(i,j)) + 1) B_{n_1,i}(\theta_1) B_{n_2,j}(\theta_2)\Big)$$

$$= \gamma M' > M$$

and

$$\min_{\theta_2 \geq \theta_1 + d} Q(\delta^{\gamma}) = \gamma M + (1-\gamma) > M \quad (\text{since } M < 1)$$

Contradicting the assumption that $\delta_0 \in \mathcal{D}_0^{n_1, n_2}$. Hence the result follows

(ii) Let $\epsilon > 0$ and suppose $\delta_0(i_0+1, j_0-1) < 1$. Let

$$\eta_{i_o,j_o} = -\varepsilon \left\{ \binom{n_1}{i_o} \binom{n_2}{j_o} \right\}^{-1}$$

and

$$\eta_{i_o+1,j_o-1} = -\varepsilon \left\{ \binom{n_1}{i_o+1} \binom{n_2}{j_o-1} \right\}^{-1}.$$

Now consider the decision rule δ^* defined as

$$\begin{aligned}
\delta^*(i,j) &= \delta_o(i,j) && \text{for } (i,j) \neq (i_o,j_o),(i_o+1,j_o-1) \\
&= \delta_o(i,j)+\eta_{i_o,j_o} && \text{for } (i,j) = (i_o,j_o) \\
&= \delta_o(i,j)+\eta_{i_o+1,j_o-1} && \text{for } (i,j) = (i_o+1,j_o-1)
\end{aligned}$$

where $\varepsilon > 0$ is chosen such that

$0 \le \delta^*(i_o,j_o) \le 1$ and $0 \le \delta^*(i_o+1,j_o-1) \le 1$.

Then

$$\begin{aligned}
Q(\delta^*) &= Q(\delta_o) + \varepsilon\, \theta_1^{i_o+1} (1-\theta_1)^{n_1-i_o-1} \theta_2^{j_o-1} (1-\theta_2)^{n_2-j_o+1} \\
&\quad - \varepsilon\, \theta_1^{i_o} (1-\theta_1)^{n_1-i_o} \theta_2^{j_o} (1-\theta_2)^{n_2-j_o} \\
&= Q(\delta_o) + \varepsilon\, \theta_1^{i_o} (1-\theta_1)^{n_1-i_o} \theta_2^{j_o-1} (1-\theta_2)^{n_2-j_o} \\
&\qquad (\theta_1(1-\theta_2)-(1-\theta_1)\theta_2) \\
&= Q(\delta_o) + \varepsilon\, \theta_1^{i_o} (1-\theta_1)^{n_1-i_o-1} \theta_2^{j_o-1} (1-\theta_2)^{n_2-j_o} (\theta_1-\theta_2)
\end{aligned}$$

Clearly,

$$Q(\delta^*) > Q(\delta_o) \qquad \text{if } \theta_1 \ge \theta_2+d$$

and $\quad Q(\delta^*) < Q(\delta_o) \qquad$ if $\theta_2 \ge \theta_1+d$

Therefore,

$$\min_{\theta_1 \ge \theta_2+d} Q(\delta^*) > \min_{\theta_2 \ge \theta_1+d} Q(\delta_o)$$

and

$$\min_{\theta_2 \ge \theta_1 + d} (1-Q(\delta^*)) > \min_{\theta_2 \ge \theta_1 + d} (1-Q(\delta_o))$$

which contradicts the assumption that $\delta_o \in \mathcal{D}_o^{n_1,n_2}$. Hence the result follows.

(iii) Note that d > .5 implies that

$$\left.\begin{array}{ll} \theta_1 > .5 & \text{for } \theta_1 \ge \theta_2 + d \\ \text{and} & \\ \theta_1 < .5 & \text{for } \theta_2 \ge \theta_1 + d \end{array}\right\} \qquad (5.3.1)$$

Suppose $\delta_o(i_o+1, j_o) < 1$ for some (i_o, j_o).

For $\varepsilon > 0$, consider the decision rule δ^* defined as :

$$\begin{aligned} \delta^*(i,j) &= \delta_o(i,j) && \text{for } (i,j) \ne (i_o,j_o), (i_o+1,j_o) \\ &= \delta_o(i,j) - \varepsilon \left\{\binom{n_1}{i_o}\binom{n_2}{j_o}\right\}^{-1} && \text{for } (i,j) = (i_o,j_o) \\ &= \delta_o(i,j) + \varepsilon \left\{\binom{n_1}{i_o+1}\binom{n_2}{j_o}\right\}^{-1} && \text{for } (i,j) = (i_o+1,j_o) \end{aligned}$$

where $\varepsilon > 0$ is chosen such that δ^* is a valid decision rule.

$$\begin{aligned} Q(\delta^*) &= Q(\delta_o) + \varepsilon\, \theta_1^{i_o+1}(1-\theta_1)^{n_1-i_o-1}\, \theta_2^{j_o}(1-\theta_2)^{n_2-j_o} \\ &\quad - \varepsilon\, \theta_1^{i_o}(1-\theta_1)^{n_1-i_o}\, \theta_2^{j_o}(1-\theta_2)^{n_2-j_o} \\ &= Q(\delta_o) + \varepsilon\, \theta_1^{i_o}(1-\theta_1)^{n_1-i_o-1}\, \theta_2^{j_o}(1-\theta_2)^{n_2-j_o}(2\theta_1-1) \end{aligned}$$

using (5.3.1), it follows that

$$Q(\delta^*) > Q(\delta_o) \quad \text{if} \quad \theta_1 \ge \theta_2 + d$$

and

$$Q(\delta^*) < Q(\delta_o) \quad \text{if} \quad \theta_2 \ge \theta_1 + d$$

Hence,

$$\min_{\theta_1 \ge \theta_2 + d} Q(\delta^*) > \min_{\theta_1 \ge \theta_2 + d} Q(\delta_o)$$

and
$$\min_{\theta_2 \ge \theta_1 + d} (1 - Q(\delta^*)) > \min_{\theta_2 \ge \theta_1 + d} (1 - Q(\delta_o))$$

which contradicts the assumption that $\delta_o \in \mathcal{D}_o^{n_1, n_2}$. Thus

$$\delta_o(i_o + 1, j_o) = 1$$

$\delta_o(i_o, j_o) > 0 \Longrightarrow \delta_o(i_o, j_o - 1) = 1$ follows similarly. ■

As a consequence of the above Theorem, we have

Corollary 5.3.2 : (i) Suppose that $\delta_o \in \mathcal{D}_1^{n_1, n_2}$ and suppose n_1 and n_2 are both positive odd integers. Then $\delta_o \in \mathcal{D}_o^{n_1, n_2}$ implies that

$$\left.\begin{aligned} &\delta_o\left(\frac{n_1+1}{2} + i, \frac{n_2-1}{2} - i\right) = 1 \quad \text{for } i \ge 0 \\ &\text{and} \\ &\delta_o\left(\frac{n_1-1}{2} - i, \frac{n_2+1}{2} + j\right) = 0 \quad \text{for } i \ge 0 \end{aligned}\right\} \qquad (5.3.2)$$

(ii) Suppose that $d > .5$ and $\delta_o \in \mathcal{D}_1^{n_1, n_2}$. Let n_1 and n_2 be both positive odd integers. Then $\delta_o \in \mathcal{D}_o^{n_1, n_2}$ implies that

$$\left.\begin{aligned} &\delta_o\left(\frac{n_1+1}{2} + i, \frac{n_2-1}{2} - j\right) = 1 \qquad \forall\ i \ge 0,\ j \ge 0 \\ &\text{and} \\ &\delta_o\left(\frac{n_1-1}{2} - i, \frac{n_2+1}{2} + j\right) = 0 \qquad \forall\ i \ge 0,\ j \ge 0. \end{aligned}\right\} \qquad (5.3.3)$$

(iii) Suppose that $d > .5$ and $\delta_o \in \mathcal{D}_1^{n_1,n_2}$. Let n_1 be odd and n_2 be even. Then $\delta_o \in \mathcal{D}_o^{n_1,n_2}$ implies that

$$\left.\begin{aligned} &\delta_o\left(\frac{n_1+1}{2} + i,\ \frac{n_2}{2} - j\right) = 1 \qquad \forall\ i \geq 0,\ j \geq 0 \\ &\delta_o\left(\frac{n_1-1}{2} - i,\ \frac{n_2}{2} + j\right) = 0 \qquad \forall\ i \geq 0,\ j \geq 0. \end{aligned}\right\} \tag{5.3.4}$$

(iv) Suppose that $d > .5$ and $\delta_o \in \mathcal{D}_1^{n_1,n_2}$. Let n_1 and n_2 be both positive even integers. Then

$$\left.\begin{aligned} &\delta_o\left(\frac{n_1}{2},\ \frac{n_2}{2}\right) = .5 \\ &\delta_o\left(\frac{n_1}{2} + i,\ \frac{n_2}{2}\right) = 1 \qquad \forall\ i \geq 1 \\ &\delta_o\left(\frac{n_1}{2},\ \frac{n_2}{2} - j\right) = 1 \qquad \forall\ j \geq 1 \\ &\delta_o\left(\frac{n_1}{2} + i,\ \frac{n_2}{2} - j\right) = 1 \qquad \forall\ i \geq 1,\ j \geq 1 \\ &\delta_o\left(\frac{n_1}{2} - i,\ \frac{n_2}{2}\right) = 0 \qquad \forall\ i \geq 1 \\ &\delta_o\left(\frac{n_1}{2},\ \frac{n_2}{2} + j\right) = 1 \qquad \forall\ j \geq 1 \\ &\text{and} \\ &\delta_o\left(\frac{n_1}{2} - i,\ \frac{n_2}{2} + j\right) = 0 \qquad \forall\ i \geq 1,\ j \geq 1 \end{aligned}\right\} \tag{5.3.5}$$

(v) Suppose that $d > .5$ and $\delta_o \in \mathcal{D}_o^{n_1,n_2}$. Then

$$\left.\begin{aligned} &\min_{\theta_1 \geq \theta_2 + d} Q(\delta_o) = \min_{\theta_1 = \theta_2 + d} Q(\delta_o)) \\ &\text{and} \\ &\min_{\theta_2 \geq \theta_1 + d} (1 - Q(\delta^*)) > \min_{\theta_2 = \theta_1 + d} (1 - Q(\delta_o)) \end{aligned}\right\}$$

Proof : (i) Suppose

$$\delta_o\left[\frac{n_1+1}{2}, \frac{n_2-1}{2}\right] = \varepsilon, \; 0 \le \varepsilon \le 1$$

Then

$$\delta_o\left[\frac{n_1-1}{2}, \frac{n_2+1}{2}\right] = 1-\delta_o\left[n_1 - \frac{n_1-1}{2}, n_2 - \frac{n_2+1}{2}\right]$$

$$= 1-\delta_o\left[\frac{n_1+1}{2}, \frac{n_2-1}{2}\right]$$

$$= 1-\varepsilon.$$

Now $\varepsilon < 1$, implies that

$$\delta_o\left[\frac{n_1+1}{2}, \frac{n_2-1}{2}\right] < 1,$$

and

$$\delta_o\left[\frac{n_1-1}{2}, \frac{n_2+1}{2}\right] > 0$$

which contradicts (ii) of Theorem 5.3.1. Thus

$$\delta_o\left[\frac{n_1+1}{2}, \frac{n_2-1}{2}\right] = 1$$

and

$$\delta_o\left[\frac{n_1-1}{2}, \frac{n_2+1}{2}\right] = 0$$

Now using (ii) of Theorem 5.3.1, we get

$$\delta_o\left[\frac{n_1+1}{2} + i, \frac{n_2-1}{2} - i\right] = 1 \quad \text{for } i \ge 0$$

$$\delta_o\left[\frac{n_1-1}{2} - i, \frac{n_2+1}{2} + i\right] = 0 \quad \text{for } i \ge 0$$

(ii) Proceeding as for the proof (i), we have

$$\delta_o\left[\frac{n_1+1}{2}, \frac{n_2-1}{2}\right] = 1$$

and

$$\delta_o\left[\frac{n_1-1}{2}, \frac{n_2+1}{2}\right] = 0$$

Now on using (iii) of Theorem 5.3.1, we get the required result.

(iii) Suppose

$$\delta_o\left(\frac{n_1+1}{2} , \frac{n_2}{2} \right) = \varepsilon, \qquad 0 \le \varepsilon \le 1$$

Then

$$\delta_o\left(\frac{n_1-1}{2} , \frac{n_2}{2} \right) = 1 - \delta_o\left(n_1 - \frac{n_1-1}{2} , n_2 - \frac{n_2}{2} \right)$$

$$= 1 - \delta_o\left(\frac{n_1+1}{2} , \frac{n_2}{2} \right)$$

$$= 1-\varepsilon$$

$\varepsilon < 1$, implies that

$$\delta_o\left(\frac{n_1+1}{2} , \frac{n_2}{2} \right) < 1,$$

and

$$\delta_o\left(\frac{n_1-1}{2} , \frac{n_2}{2} \right) > 0$$

which contradicts (iii) of Theorem 5.3.1. Thus,

$$\delta_o\left(\frac{n_1+1}{2} , \frac{n_2}{2} \right) = 1$$

and

$$\delta_o\left(\frac{n_1-1}{2} , \frac{n_2}{2} \right) = 0$$

Result follows from (iii) of Theorem 5.3.1

(iv) $$\delta_o\left(\frac{n_1}{2} , \frac{n_2}{2} \right) = 1 - \delta_o\left(n_1 - \frac{n_1}{2} , n_2 - \frac{n_2}{2} \right)$$

$$\longrightarrow \delta_o\left(\frac{n_1}{2} , \frac{n_2}{2} \right) = .5$$

and result follows from (iii) of Theorem 5.3.1.

(v) Using (iii) of Theorem 5.3.1 and Lemma 3.2 of Risko we get the required result.■

Remark : Using the fact that $\mathcal{D}_1^{n_1,n_2}$ is an essentially complete class, we conclude that there exist a minimax rule satisfying (5.3.2) - (5.3.5).

Note that the class of rules considered by Risko

$$\mathcal{D}_R^{n_1,n_2} = \left\{ \delta_{\beta_1,\beta_2} : \delta_{\beta_1,\beta_2}(i,j) = \psi\{\beta_1+(1-2\beta_1+\beta_2)\frac{i}{n_1} - \beta_2\frac{j}{n_2}\} \right\}$$
$$(\beta_1,\beta_2) \in T = \{(x,y) \in \mathbb{R}^2 | \ x \le \tfrac{1}{2}(1+y),\ y \ge 0\},$$

where

$$\psi(x) = \begin{cases} 1 & \text{if } x \ge 1 \\ x & \text{if } 0 < x < 1, \\ 0 & \text{if } x \le 0 \end{cases}$$

satisfies (i) of Theorem 5.3.1. However, the following examples shows that rules in $\mathcal{D}_R^{n_1,n_2}$ in general do not satisfy (ii) and (iii) of Theorem 5.3.1.

Example 5.3.1 : Let $n_1 = 5$, $n_2 = 4$, $\beta_1 = \beta_2 = 0.5$. Then

$$\delta_{\beta_1,\beta_2}(0,1) = \frac{1}{2} - \frac{1}{8} = \frac{3}{8}$$

whereas,

$$\delta_{\beta_1,\beta_2}(1,0) = \frac{1}{2} + \frac{1}{10} = \frac{3}{5} < 1.$$

Thus δ_{β_1,β_2} does not satisfy (ii) of Theorem 5.3.1.

Example 5.3.2 : Choose $n_1 = 4$, $n_2 = 10$, $\beta_1 = 1$, $\beta_2 = 3$. Then

$$\delta_{\beta_1,\beta_2}(3,8) = 0.1$$

and

$$\delta_{\beta_1,\beta_2}(4,7) = 0.9 < 1.$$

Hence (ii) of Theorem 5.3.1 is not satisfied.

The following two examples establish that decision rules in Risko's class in general do not satisfy (iii) of Theorem 5.3.1.

Example 5.3.3 : Suppose $n_1 = 15$, $n_2 = 20$, $\beta_1 = -1$, $\beta_2 = 5$. Then

$$\delta_{\beta_1,\beta_2}(3,1) = .35$$

and

$$\delta_{\beta_1,\beta_2}(4,1) = \frac{32}{15} - 1.25 < 1.$$

Thus, in this case (iii) of Theorem 5.3.1 is not satisfied.

Example 5.3.4 : Let $n_1 = 2$, $n_2 = 4$, $\beta_1 = 2$, $\beta_2 = 3.5$. Then

$$\delta_{\beta_1,\beta_2}(1,2) = 0.5$$

and

$$\delta_{\beta_1,\beta_2}(2,2) = 0.85 < 1.$$

and hence (iii) of Theorem 5.3.1 is not satisfied.

REFERENCES

Alam,K. (1970) : "A two-sample procedure for selecting the population with the largest mean from k normal populations," Annals of the Institute of Statistical Mathematics, 22, 127-136.

Alam,K. (1973) : "On a multiple decision rule," Annals of Mathematical Statistics, 1, 750-755.

Alam,K. and Rizvi,M.H. (1966) : "Selection from multivariate populations," Annals of the Institute of Statistical Mathematics, 18, 307-318.

Ambughalous,M.M. and Miescke,K.J. (1989) : "On selecting the largest success probability with unequal sizes," Journal of Statistical Planning and Inference, 21, 53-68.

Bahadur,R.R. (1950) : "On ~~the~~ a problem in the theory of k populations," Annals of Mathematical Statistics, 21, 362-375.

Bahadur,R.R. and Goodman,L.A. (1952) : " Impartial decision rules and sufficient statistics," Annals of Mathematical Statistics, 23, 533-562.

Bahadur,R.R. and Robbins,H. (1950) : "The problem of greater mean," Annals of Mathematical Statistics, 21, 469-487.

Bechhofer,R.E. (1954) "A single sample multiple-decision procedure for ranking means of normal normal populations," Annals of Mathematical Statistics, 25, 16-39.

Bechhofer,R.E., Dunnett,C.W., and Sobel,M. (1954): "A two sample multiple decision procedure for ranking means of normal populations with a common unknown variance," Biometrika, 41, 170-176.

Berger,J.O. (1985) : Statistical Decision theory and Bayesian Analysis, Springer-Verlag, New York.

Berger,R.L. (1977) : "Minimax, admissible, and gamma minimax multiple decision rules," Technical Report No. 489, Department of Statistics, Purdue University, Indiana.

Berger,R.L. (1979) : "Minimax subset selection for the loss measured by subset size," Annals of Statistics, 7, 1333-1338.

Berger,R.L. (1980) : "Minimax subset selection for the multinomial distribution," Journal of Statistical Planning and Inference, **4**, 391-402.

Berger,R.L. (1982) : "A minimax and admissible subset rule for the least probable multinomial cell," Statistical Decision Theory and Related Topics III, Volume 1, (Edited by S.S. Gupta and J.O. Berger), Academic press, New york, 143-156.

Berger,R.L. and Gupta,S.S. (1980) : "Minimax subset selection rules with applications to unequal variances (unequal sample sizes) problems," Scandvanian Journal of Statistics, **7**, 21-26.

Bhandari,S.K. and Chaudhuri,A.R. (1990) : "On two conjectures about two-stage selection procedures," Sankhya, Series A, **52**, 131-141.

Bickel,P.J. and Yahav,J.A. (1977) : "On selecting a subset of good populations," Statistical Decision Theory and Related Topics II, (Edited by S.S. Gupta and D.S. Moore), Academic Press New york, 37-55.

Bjornstad, J.F. (1978): "On optimal subset selection procedures," Ph.D. Dissertation Department of Statistics, Univ. of Calif., Berkeley.

Bjornstad,J.F. (1981a) : "A decision-theoretic approach to subset selection," Communications in Statistics-Theory and Methods, **10**, 2411-2433.

Bjornstad,J.F. (1981b) : "A class of Schur procedures and minimax theory for subset selection," Annals of Statistics, **9**, 777-791.

Bjornstad,J.F. (1984) : "A general theory of asymptotic consistency for subset selection with applications," Annals of Statistics, **12**, 1058-1070.

Bjornstad,J.F. (1985) : Asymptotic consistency of procedures for selecting good populations," Communications in Statistics-Theory and Methods, **14**, 1659-1668.

Bjornstad,J.F. (1986) : "Asymptotic consistency for subset selection procedures satisfying the P^*-condition," Journal of Statistical Planning and Statistical Inference, **13**, 319-335.

Blackwell,D. and Girshick,M.A. (1954) : Theory of Games and Statistical Decisions, John Wiley and Sons, New york.

Bland, R.P.and Bratcher,T.L. (1968) : "A Bayesian approach to the problem of ranking binomial populations," Siam Journal of Applied Mathematics, **16**, 843-850.

Blum~~m~~,J.R. and Rosenblatt, J. (1967) : "On partial apriori information in statistical inference," Annals of Mathematical Statistics, **38**, 1671-1678.

Chernoff,H.,and Yahav,J.A. (1977) : "A subset selection problem employing new criterion," Statistical Decision Theory and Related Topics II, (Edited by S.S.Gupta and D.S.Moore), Academic press, New York, 93-119.

Deely,J.J. (1965) : "Multiple decision procedures from an empirical Bayes approach," Mimeograph Series **No. 45**, Department of Statistics, Purdue University, Indiana.

Deely,J.J. and Gupta,S.S. (1968) : "On the properties of subset selection procedures," Sankhya Series A, **30**, 37-50.

Desu,M.M. and Sobel,M. (1968) : "A fixed subset size approach to selection problem," Biometrika, **55**, 401-410.

Dhariyal,I.D., Misra,N., and Rathore,R.K.S. (1989) : "Selecting the better of two binomial populations : Optimal decision rules, Calcutta Statistical Association Bulletin, **39**, 157-167.

Dudewicz,E.J. (1971) : "Nonexistence of a single sample selection rule whose P(CS) is independent of variances," South African Statistical Journal, **5**, 37-39.

Dudewicz,E.J. and Dalal,S.R. (1975) : "Allocation of observations in ranking and selection with unequal variances," Sankhya Series B, **37**, 28-78.

Dunnett, C.W. (1960): "On selecting the largest of k normal population means," Journal of the Royal Statistical Society, 22B, 1-40.

Eaton,M.L. (1967) : "Some optimum properties of ranking procedures," Annals of Mathematical Statistics, **38**, 124-137.

Goel,P.K. and Rubin,H. (1977) : "On selecting a subset containing the best population - A Bayesian approach," Annals of Statistics, **5**, 969-983.

Gupta,S.S. (1956) : "On a decision rule for a problem in ranking normal means," Mimeograph Series **No. 150**, University of North Carolina Chapel Hill.

Gupta,S.S. (1963) : "On a ranking and selection procedure for gamma populations," Annals of the Institute of Statistical Mathematics, **14**, 199-216.

Gupta,S.S. (1965) : "On some multiple decision (selection and ranking) rules," Technometrics, **7**, 225-245.

Gupta,S.S. (1966) : "On some selection and ranking procedures for multivariate normal populations," Multivariate Analysis, (Edited by P.R.Krishnaiah), Academic Press, New York, 457-475.

Gupta,S.S. and Hsu,J.C. (1978) : "On the performance of some subset selection procedures," Communications in Statistics - Simulation and Computation, 7, 561-591.

Gupta,S.S. and Hsiao,P. (1983) : "Empirical Bayes rules for selecting good populations," Journal of Statistical Planning and Statistical Inference, 8, 87-101.

Gupta,S.S. and Huang,D.Y. (1975) : "On subset selection procedures for Poisson populations and some applications to the multinomial selection problems," Applied Statistics, (Edited by R.P.Gupta), North-Holland, Amsterdam, 97-109.

Gupta,S.S. and Huang,D.Y. (1976) : "Subset selection procedure for the means and variances of normal populations : Unequal sample sizes case," Sankhya, Series B, **38**, 112-128.

Gupta,S.S. and Huang,D.Y. (1977) : "On some Γ-minimax procedures and multiple comparison procedures," Statistical Decision Theory and Related Topics II, (Edited by S.S.Gupta and D.S.Moore), Academic Press, New York, 139-155.

Gupta,S.S. and Huang,D.Y. (1980a) : "A note on optimal subset selection procedures," Annals of Statistics, **8**, 1164-1167.

Gupta,S.S. and Huang,D.Y. (1980b) : "An essentially complete class of multiple decision procedures," Journal of Statistical Planning and Statistical Inference, **4**, 115-121.

Gupta,S.S., Huang,D.Y., and Nagel,K. (1979) : "Locally optimal subset selection procedures based on ranks," Optimizing Methods in Statistics, (Edited by J.S.Rustagi), Academic Press, New york, 251-260.

Gupta,S.S. and Kim,W.-C. (1980) :"Γ-minimax and minimax decision rules for comparison of treatments with a control," Recent Developments in Statistical Inference and Data Analysis, (Edited by K.Matusita) North-Holland, Amsterdam, 55-71.

Gupta,S.S. and Kim,W.-C. (1984) : "A two-stage elimination type procedure for selecting the largest of several normal means," Designs of Experiments : Ranking and Selection, (Edited by T.J.Santner and A.C.Tamhane), Marcel Dekker, New York, 77-93.

Gupta,S.S. and Leu,L.Y. (1983) :"On Bayes and empirical Bayes rules for selecting good populations," Technical Report **No. 83-37**, Department of Statistics, Purdue University, Indiana.

Gupta,S.S. and Liang,T. (1984) : "Empirical Bayes rules for selecting good binomial populations," Technical Report **No. 84-37**, Department of Statistics, Purdue University, Indiana.

Gupta,S.S. and Liang,T. (1986a) : "Empirical Bayes rules for selecting the best binomial population," Technical Report No. 86-13, Department of Statistics, Purdue University, Indiana.

Gupta,S.S. and Liang,T. (1986b) : "Parametric empirical Bayes rules for selecting the most probable multinomial event," Technical Report No. 86-54, Department of Statistics, Purdue University, Indiana.

Gupta,S.S. and Mc Donald,G.C. (1970) : "On some classes of selection procedures based on ranks," Nonparametric Techniques in Statistical Inference, (Edited by M.L.Puri), Cambridge University Press, London, 491-514.

Gupta,S.S. and Miescke,K.J. (1982) : "On the least favorable configuration in certain two-stage selection procedures," Statistics and Probability : Essays in Honor Of C.R.Rao, (Edited by G.Kallianpur, P.R.Krishnaiah, and J.K.Ghosh), North-Holland, Amsterdam, 295-305.

Gupta,S.S. and Miescke,K.J. (1983) : "An essentially class of two-stage procedures with screening at first stage," Statistics and Decisions, **1**, 427-439.

Gupta,S.S. and Miescke,K.J. (1984a) : "On two-stage Bayes selection procedures," Sankhya, Series B, **46**, 123-134.

Gupta,S.S. and Miescke,K.J. (1984b) : "Sequential selection procedures-A decision theoretic approach," Annals of Statistics, **12**, 336-350.

Gupta,S.S. and Miescke,K.J. (1985) : "Minimax multiple t-tests for comparison of k normal populations with a control," Journal of Statistical Planning And Inference, **12**, 161-169.

Gupta,S.S. and Miescke,K.J. (1986) : "On the problem of finding the largest normal mean under heteroscedasticity," Statistical Decision Theory and Related Topics-IV, Volume 2, (Edited by S.S.Gupta and J.O.Berger), Springer-Verlag, Berlin, 37-49.

Gupta,S.S. and Nagel,K. (1967) : "On selection and ranking procedures and order statistics from the multinomial distribution," Sankhya, Series B, **29**, 1-34.

Gupta,S.S. and Nagel,K. (1971) : "On some contributions to multiple decision theory," Statistical Decision Theory and Related Topics, (Edited by S.S.Gupta and J.Yackel), Academic Press, New York, 79-102.

Gupta,S.S. and Panchpakesan,S. (1972) : "On a class of subset selection procedures," Annals of Mathematical Statistics, **43**, 824-822.

Gupta, S.S. and Panchpakesan, S. (1979) : Multiple Decision Procedures : Theory and Methodology of Selecting and Ranking Populations, John Wiley and Sons, New York.

Gupta, S.S. and Sobel, M. (1958) : "On selecting a subset which contains all populations better than a standard," Annals of Mathematical Statistics, 29, 235-244.

Gupta, S.S. and Sobel, M. (1960) : "selecting a subset containing the best of several binomial populations," Contributions to Probability and Statistics, Essays in Honor of Harold Hotelling, (Edited by I.Olkin, G.Ghuraye, W.Hoeffding, M.G.Maddow, and H.B.Mann), Stanford University Press, Stanford, California, 224-248.

Gupta, S.S. and Wong, W.-Y. (1982) : "Subset selection procedures for the means of normal populations with unequal variances : unequal sample sizes case," Selecta Statistica Canadiana, 6, 109-149.

Gupta, S.S. and Yang, H.-M. (1985) : "Bayes-P^* subset selection procedures for the best population," Journal of Statistical Planning and Inference, 12, 213-233.

Hall, W.J. (1959) : "The most economical character of Bechhofer and Sobel decision rules," Annals of Mathematical Statistics, 30, 964-969.

Hochberg, Y. and Marcus, R. (1981) : "Three-stage elimination type procedures for selecting the best normal population when variances are unknown," Communications in Statistics - Theory and Methods, 10, 597-612.

Hsu, T.-A. (1985) : "Some optimal subset selection procedures with applications," Sankhya, Series B, 47, 206-215.

Huang, D.-Y. and Panchpakesan, S. (1982) : "Some locally optimal subset selection procedures based on ranks," Statistical Decision Theory and Related Topics - III, Volume 2, (Edited by S.S.Gupta and J.O.Berger), Academic Press New York, 1-14.

Huang, D.-Y., Panchpakesan, S. and Tseng, S.-T. (1984) : "Some locally optimal selection rules for comparison with a control," Journal of Statistical Planning and Inference, 9, 63-72.

Lehmann, E.L. (1966) : "On a theorem of Bahadur and Goodman," Annals of Mathematical Statistics, 37, 1-6.

Miescke, K.J. (1979) : "Bayesian subset selection for additive and linear loss functions," Communications in Statistics - Theory and Methods, 8, 1205-1226.

Miescke,K.J.(1984) : "Recent results on multi-stage selection procedures," Proceedings of the Seventh Conference on Probability Theory, (Edited by M.Iosifescu), Editura Academiei Republicii Socialiste, Romania Bucharest, 259-268.

Miescke,K.J. and Sehr,J. (1980) : "On a conjecture concerning least favorable configuration of certain two-stage procedures," Communications in Statistics - Theory and Methods, 9, 1609-1617.

Mishra,S.N. and Dudewicz,E.J. (1984) : "Simultaneuos selection of extreme populations : A subset selection approach," Technical Report No. S-23, Department of Mathematics, Syracuse University, Syracuse, New York.

Mishra,S.N. (1986a) : "Simultaneous selection of extreme populations scale parameter case," Prog. of Maths, 20, 91-105.

Mishra,S.N. (1986b) : "Simultaneous selection of extreme population means : Indifference-zone formulation," American Journal of Mathematical and Management Sciences, 6, 131-142.

Mishra,S.N. (1986c) : "A nonparametric procedure for selection of extreme populations based on quantiles,". To appear in IMSIBAC.

Mishra,S.N. and Dudewicz,E.J. (1987) : "Simultaneuos selection of extreme populations : A subset selection approach," Biometrical Journal, 29, 471-483.

Mosteller,F. (1948) : "A k-sample slippage test for an extreme population," Annals of Mathematical Statistics, 19, 58-65.

Nagel,K. (1970) : "On subset selection rules with certain optimality properties," Mimeograph Series No. 222, Department of Statistics Purdue University, Indiana.

Nelder,J.A. and Mead,R. (1965) : "A simplex method for function minimization," Computer Journal, 7, 308-313.

Olsson,D.M. (1974) : "A sequential simplex program for solving minimization problems," Journal of Quality Technology, 6, 53-57.

Paulson,E. (1949) : "A multiple decision procedure for certain problems in analysis of variance," Annals of Mathematical Statistics, 20, 95-98.

Paulson,E. (1952) : "On the comparison of several experimental categories with a control," Annals of Mathematical Statistics, 23, 239-246.

Randles,R.H. and Hollander,M. (1971) : "Γ-minimax selection procedures in treatment versus control problems," Annals of Mathematical Statistics, 42, 330-341.

Risko,K.J. (1985) : "Selecting the better binomial population with unequal sample sizes," Communications in Statistics - Theory and Methods, **14**, 123-158.

Robbins,H. (1955) : "An empirical Bayes approach to statistics," Proceedings of Third Berkeley Symposium on Mathematical Statistics and Probability, (Edited by J.Neyman), University of California Press, Berkeley and Los Angeles, 157-163.

Robbins,H. (1964) : "The empirical Bayes approach to statistical decision problems," Annals of Mathematical Statistics, **35**, 1-19.

Seal,K.C. (1955) "On a class of decision procedures for ranking means of normal populations," Annals of Mathematical Statistics, **26**, 387-398.

Seal,K.C. (1957) : "An optimum decision rule for ranking means of normal populations," Calcutta Statistical Association Bulletin, **7**, 131-150.

Seal,K.C. (1958) : "On ranking parameters of scale in Type-III populations," Journal of the American Statistical Association, **53**, 164-175.

Sobel,M. and Huyett,M.J. (1957) : "Selecting the best one of several binomial populations," Bell System Technical Journal, **36**, 537-576.

Studden,W.J. (1967) : "On selecting a subset of k populations containing the best," Annals of Mathematical Statistics, **38**, 1072-1078.

Tamhane,A.C. (1976) "A three-stage elimination type procedure for selecting the largest mean (Common unknown variance)," Sankhya, Series B, **38**, 339-349.

Tamhane,A.C. and Bechhofer,R.E. (1977) : "A two-stage minimax procedure with screening for selecting the largest normal mean," Communications in Statistics - Theory and Methods, **6**, 1003-1033.

Tamhane,A.C. and Bechhofer,R.E. (1979) : "A two-stage minimax procedure with screening for selecting the largest normal mean," (II) : An improved PCS lower bound and associated tables, Communications in Statistics - Theory and Methods, **8**, 337-358.

Van Ryzin,J. (1970) : "On some nonparametric empirical Bayes multiple decision problems," Proceedings of First International Symposium on Nonparametric Techniques in Statistical Inference, (Edited by M.L.Puri), 585-603, Cambridge University Press, London.

Date Slip

This book is to be returned on the date last stamped.
